41

ajer

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XLI NUMBER 1 MARCH 1995

The Alberta Journal of Educational Research

Published in March, June, September, and December by the Faculty of Education, University of Alberta.

AJER is a quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Judy Cameron

Managing Editor: Naomi Stinson
Proofreader: Karen McFarlane

Consulting Editors

Terry Belke
Mount Allison University

George Buck
University of Alberta

Ardra L. Cole
Ontario Institute for Studies in Education

John Connors
Canadian Union College

Samuel Deitz
Georgia State University

Sharon M. Haggerty
University of Western Ontario

Antoinette Oberg
University of Victoria

Ruth Rees
Queen's University

Heather Ryan
Massey University

Robert H. Short
University of Alberta

Kelleen Toohey
Simon Fraser University

Robert Wilson
Queen's University

Faculty Advisory Committee

John G. Paterson (Chair)
Joyce Edwards
Ken Ward
Robert H. Short

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, the Alberta Advisory Committee for Educational Studies, and the Research Grants Office of the University of Alberta

The subscription rate is $32.00 per year for individuals, $45.00 per year for institutions. Add $8.00 for delivery outside Canada. Single copies are $8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to:

The Alberta Journal of Educational Research
Office of the Dean
845 Education Centre South
University of Alberta
Edmonton, AB, Canada, T6G 2G5.
Fax: (403) 492-0390

Publications Mail Registration Number 1436

Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XLI, Number 1, March 1995

Articles

John Biggs **1** Assessing for Learning: Some Dimensions Underlying New Approaches to Educational Assessment

Debora Barnett-Foster
Philip Nagy **18** A Comparison of Undergraduate Test Response Strategies for Multiple-choice and Constructed-response Questions

Jeremy Hull
Ron Phillips
Eleoussa Polyzoi **36** Indian Control and the Delivery of Special Education Services to Students in Band-operated Schools in Manitoba

Grace V. Malicky
Charles A. Norman **63** Perceptions of Literacy and Adult Literacy Programs

Richard Conte
Jac J.W. Andrews
Melanie Loomer
Gillian Hutton **84** A Classroom-based Social Skills Intervention for Children with Learning Disabilities

John A. Ross
J. Bradley Cousins **103** Giving and Receiving Explanations in Cooperative Learning Groups

Book Reviews

Tracey Derwing **122** *Genie: An Abused Child's Flight from Silence* by Russ Rymer

Gregor Wolbring **124** *Violence and Abuse in the Lives of People With Disabilities* by Dick Sobsey

ISSN 0002-4805

Acknowledgment

The following people have helped to maintain the quality of AJER by reviewing manuscripts during the past year. Their contributions are greatly appreciated.

Assheton-Smith, M.
Bain, B.
Baine, D.
Bateson, D.
Beebe, M.
Belke, T.
Benson, G.
Bilash, O.
Blakey, J.
Carson, T.
Chambers, C.
Chapman, J.
Cole, A.
Connors, J.
Dawson, S.
Derwing, T.M.
Donald, J.
Draper, J.
Edwards, J.
Epling, W.F.
Gaskell, J.
Glegg, A.
Goldberg, J.
Golic, J.
Haggerty, S.
Harley, B.
Housego, B.E.J.
Juliebo, J.
Kalish, C.
Kirby, J.
Kirman, J.
Konrad, A.
Krentz, C.
Krull, C.
Kysela, G.M.
LaRocque, L.
Leithwood, K.
LeRoy, C.
Maguire, M.
Maguire, T.O.
Marini, Z.
McCann, S.
McFadden, C.
Mohan, B.
Norris, J.
Norris, S.
Oberg, A.
Owens, D.T.
Pearson, A.
Phillips, L.
Pierce, W.D.
Piper, T.
Ray, D.
Rees, R.
Rogers, T.
Rowell, P.
Ryan, B.
Ryan, H.
Samiroden, W.
Snart, F.
Sobsey, R.J.
Stapleton, J.
Symons, F.
Taafe, R.
Taylor, G.
Thomas, A.
Toohey, K.
Uhl, N.
VanBrunschot, E.
Ward, K.
Wiener, J.
Wilgosh, L.
Wilson, R.J.
Winne, P.
Wyrostok, N.

The Alberta Journal of Educational Research *Vol. XLI, No. 1, 1995, 1-17*

John Biggs
University of Hong Kong

Assessing for Learning: Some Dimensions Underlying New Approaches to Educational Assessment

The theory and practice of assessing learning are currently undergoing a paradigm shift. The critical realization in producing this change is that educational considerations should drive testing, not psychometric or political ones. Three dimensions interact to yield different modes of assessment, including different kinds of performance assessment: the measurement versus the standards model of testing, quantitative and qualitative assumptions as to the nature of what is learned, and whether the learning and testing is situated or decontextualized. The modes of assessment so generated are suited for different educational aims, but the most appropriate modes are underrepresented in current practice, quantitative and decontextualized modes being greatly overrepresented, resulting in backwash often deleterious to teaching and learning. Conceptual and structural difficulties in implementing qualitative and situated modes of assessment are discussed.

La théorie et la pratique de l'évaluation de l'apprentissage subissent présentement un changement de forme draconien. Ce changement de la perspective de l'évaluation de l'apprentissage relèverait du fait que se sont les considérations pédagogiques éducationnelles qui devraient diriger l'orientation du testing et non les orientations psychométriques ou politiques. Trois dimensions s'entrecroisent produisant différents modes d'évaluations incluant différentes sortes d'habiletés et performances: le modèle de testing de la mesure comparé au modèle de testing des standards, les prédispositions qualitatives et quantitatives lorsqu'elles concernent la nature de ce qui est appris, et si l'apprentissage et le testing sont situés ou décontextualisés. Les façons d'évaluation qui en sont par conséquent créées semblent être adaptées à différents buts éducationnels, mais semblerait-il que les méthodes les plus appropriées sont sous-représentées dans la pratique courante de l'enseignement, et que les méthodes quantitatives et décontextualisées semblent être sur-représentées ce qui semble effectuer un certain recul sur l'enseignement et l'apprentissage. On discute des difficultés conceptuelles et structurelles de l'implantation des méthodes d'évaluation qualitatives et situées.

Assumptions Underlying Assessing and Learning

The Measurement Model

For many years, perhaps the greater part of this century, the assessment of academic learning proceeded with little change and with little perceived need for change. Test developers, psychometricians, and teacher educators accepted the *measurement* model of assessment, based on trait theory, which requires that test item scores are sufficiently consistent with each other that testees may be ordered along a single dimension represented by that intra-test consistency (Taylor, 1994). The technologies of test construction, item selection, and establishing reliability and validity followed from these basic assumptions. That

---

John Biggs is a professor of education in the University of Hong Kong, a post he has held since 1987. His major research area is learning and motivation in institutional settings, and he is particularly interested in the effects of approaches to teaching and assessment on the quality of student learning.

technology became sophisticated and powerful in its appropriate applications. Classroom teachers for their part probably held the accepted framework as espoused theory, their theory-in-use an uneasy compromise between that and established practice (but see Cizek, 1993); when they do try to put their espoused theory directly into practice they usually get it wrong (Marso & Pigge, 1991).

The problem is not teacher incompetence, but the fact that the "technology of assessment that grew out of test theory ... lacked a basis in psychological theory" (Wilson & Kirby, 1994, p. 107); even more to the point, it lacked a basis in educational theory and the knowledge base of teaching (Haertel, 1991). With changing educational philosophies traditional test theory became decreasingly appropriate to the majority of classroom testing occasions. Specifically, the model assumes the stability of the dimension being tested, which is appropriate when the task is to discriminate between individual performances and to predict future performance, but is inappropriate to assess individual or group outcomes in relation to a particular curriculum because the intervention of *teaching* assumes change. As social expectations became increasingly that most students should complete grades K through 12, acquiring basic skills, competences and certain declarative knowledge on the way, selecting students for different levels of schooling ceased to be a major issue, but ensuring that students attained these standards acceptably (however one defines that in the event) was a definite issue of public concern. In an age of electronic learning assessment practices were being driven by steam.

The Standards Model

Models of assessment based on outcomes reaching a predetermined standard have also been around for many years—for example, the notion of a First Class Honors in the British University system—but these ideas were subjective and they specified no technology. The criterion-referenced testing (CRT) movement was probably the first systematic attempt to formulate and enact a set of assumptions critically different from those underlying the measurement model in that instructional and assessment strategy were specifically linked (Bloom, Hasting, & Madaus, 1971; Popham & Husek, 1969). CRT is an example of what Taylor (1994) calls the standards model of assessment, which is based on such assumptions as: public standards can be set; they can be reached by most students, albeit by different kinds of performance; and fair and consistent judgments are possible to determine whether the standards have been met. These assumptions about assessment are isomorphic to those underlying learning and instruction.

The traditional CRT model fell short, however, on the nature of the standards that were set, which in some important respects were no improvement on those underlying the measurement model itself. Although CRT was an improvement in terms of the link between assessment and instruction, the links between these and the question of what is learned remained unchanged. CRT had as limited a view of learning as had the measurement model.

Two basic conceptions of the nature of learning exist in our educational thinking, quantitative and qualitative (Cole, 1990; Marton, dall'Alba, & Beaty, in press). How we view learning will determine how we go about learning, teaching, and, to the present point, assessment.

The Quantitative Tradition

The quantitative tradition has the longest history in educational thinking, stemming from the positivist tradition in the social sciences (Moss, 1992). Learning is conceived as acquiring "specific discrete skills described as precise well-delimited behaviors" (Cole, 1990, p. 2). These contents of learning are treated as discrete quanta of declarative or procedural knowledge; as far as assessment is concerned, any one quantum is treated as functionally independent of any other. In this view, the curriculum becomes in effect a list of discrete units: facts, skills, competences, behavioral objectives, performance indicators, and the like, and assessment a matter of how many of these have been attained.

Teaching is conceived as transmitting knowledge from teacher to learner, and many delivery and assessment systems are based on the transmission model up to college and university level (Trigwell, Prosser, & Taylor, 1994). The teacher's task is to know the subject and expound it clearly, the learner's to receive accurately.

In assessment practice, the contents of knowledge are seen as learned in binary units (correct/incorrect), the correct units being summed to give an aggregate score that yields an index of competence in what is learned. Multiple-choice tests enact this clearly; competence is represented as a total score of all items correct, any one item being worth the same as any other. Lohman (1993) cites an example of a multiple-choice test given to grade 5 children when the 200th anniversary of the United States Constitution was being celebrated. The only item on the test referring to Thomas Jefferson was "Who was the signer of the constitution who had six children?" The problem is that this tells the child that every idea in the test is equally important. Lohman (1993) recounts that a year later he asked a child in this class what she remembered of Thomas Jefferson: the number of his children, but nothing of his role in the Constitution. Such testing tells students that

> There is no need to separate main ideas from details; all are worth one point. And there is no need to assemble these ideas into a coherent summary or to integrate them with anything else because that is not required. (p. 19)

However, particularly in large classes where marking schemes are used, essays are frequently treated in the same manner, a mark being given as each "correct" or "acceptable" point is made, with possibly bonus points for argument or style. I am not saying that individual items in a multiple-choice test may not be clever, demanding, or substantively significant or that the essay question is unsuccessful in eliciting high cognitive-level engagement, but that the treatment of the item scores and test marks assumes their mutual equivalence, independence, and additivity. Students know this and are strategic in exploiting it (Biggs, 1973; Crooks, 1988); in timed examinations, for instance, in view of the law of diminishing returns with respect to mark allocation (time spent on the first half of an essay almost always nets more marks than the same time spent on the second half), attempting all five, say, questions but finishing none will capture more marks than writing a properly structured answer to only four. It is interesting to note that the national model in the United States for setting both objective and essay questions, the Bloom

taxonomy, "has no category for the organization of factual knowledge" (Lohman, 1993, p. 22).

Nevertheless, whether the backwash effects of multiple-choice and other testing formats in the quantitative tradition are seen as deleterious or beneficial for instruction is a much debated question, the answer no doubt depending on whether one holds a quantitative view of learning or not. There have been several warnings of the harmful influences, ranging from the popular (Hoffman, 1962) to the scholarly (Frederiksen, 1984), whereas psychometricians who hold avowed quantitative assumptions about the nature of learning see deliberate and systematic teaching to the test as sound teaching (Shepard, 1991), elevated by Popham (1987) to the strategy of measurement-driven instruction.

It will be noted that quantitative assumptions underlie both norm-referenced testing (NRT) in the traditional measurement model and criterion-referenced testing (CRT) in the Popham and mastery models. The issue is thus not simply the measurement versus the standards models (Taylor, 1994), but a question of one's assumptions about the nature of the learning to be assessed.

The Qualitative Tradition

The qualitative tradition has its roots in 19th-century phenomenology, and later in Gestalt psychology, but it is only recently that it has got as far as offering an alternative paradigm as far as educational decision making on a wide front is concerned. The underlying theory of learning is *constructivism*, a family of theories rather than any one, according to which students are assumed to learn cumulatively, actively interpreting and incorporating new material with what they already know. Different theories variously emphasize the individual, social, cognitive, saccadic, contextual, or emergent natures of learning, but all agree on an active learner seeking meaning by constructing knowledge rather than by receiving and storing knowledge. Poststructuralists emphasize the social construction of knowledge (Delandsheere & Petrovsky, 1994), whereas my own view and that underlying this article stems ultimately from cognitive psychology (Biggs & Moore, 1993). In this latter view understanding changes progressively as people learn, with qualitative changes taking place in the nature both of what is learned and how it is structured. Understanding of a topic thus evolves cumulatively over the long haul, having *horizontal* interconnections with other topics and subjects and *vertical* interconnections with previous and subsequent learnings in the same topic.

As the contents of learning are meanings, the curriculum question is to decide what meanings or levels of understanding are reasonable at the stage of learning in question. The teacher's task is not then to transmit correct understandings, but to help students construct understandings that are progressively more mature and congruent with accepted thinking, recognizing that in many subject areas students' everyday experiences have helped them to construct alternative ways of construing their world. Teaching techniques may be expository at times, but essentially they will involve more effective ways of eliciting constructive activity on the part of the student, such as the deliberate and explicit use of the relevant knowledge base, peer and student-teacher interaction, a motivating context, and much student activity, both reflective or self-directed as well as task-directed (Biggs & Moore, 1993).

Whereas the logic of assessment from a quantitative point of view implies aggregating units of learning taken cross-sectionally with respect to time, that from the qualitative tradition implies charting longitudinal growth over time, from relative ignorance to relative competence: establishing the limits of *relative* is of course a major curriculum question. The outcomes of learning become the constructions the learner has made at any given point in the process. If that growth in competence can be described in recognizable stages, then so much the better because these stages can then become assessment targets (Biggs & Collis, 1989; Clark, Scarino, & Brownell, 1994).

Assessment within the qualitative framework may be of two basic kinds:

1. *developmental*, the purpose of which is to discover where students are in the development of understanding or competence in the domain of the concept or skill in question, the focus here being pure or discipline-based knowledge;
2. *ecological*, the purpose of which is to discover if students can carry out tasks that are "worthwhile, significant, and meaningful" (Archbald & Newman, 1988, p. 1), the focus here being on applications and problem solving.

As far as issues of assessment are concerned, there is some similarity between quantitative/qualitative distinction and Lohman's distinction between crystallized abilities as being involved in teaching and assessing for near transfer in familiar situations, and likewise for fluid abilities for far transfer in unfamiliar situations; the first involve lower-order, and the second, higher-order cognitive skills. Education should involve both, but common assessment tasks evoke predominantly the former (Lohman, 1993).

Situated and Decontextualized Assessment

Ecological assessment thus becomes the qualitative assessment of applied procedural knowledge, and as such is closely related to the more general movement referred to as "authentic" assessment (Newman & Archbald, 1992; Wiggins, 1989), which insists that the context of testing should reflect the goals of learning insofar as they require students to think, decide, and act in the real world (Archbald & Newman, 1988). However, the term *authentic* grabs the moral high ground rather, and *performance* assessment is now more usual for this mode of assessment (Moss, 1992), being neutral with respect to rectitude while suggesting that the test items should require some kind of active demonstration of the knowledge in question rather than a propositional account of it. Some action is required in a realistic setting, involving enactment of a skill or problem solving. Conventional pencil-and-paper tests evidently do not meet this requirement.

Performance assessment (PA) is closely related to the concept of situated cognition. Brown, Collins, and Duguid (1989) suggest that the only valid or powerful form of learning is what takes place in situated contexts. Schools, on the other hand, assume "a separation between knowing and doing, treating knowledge as an integral, self-sufficient substance, theoretically independent of the situations in which it is learned and used. The primary concern of schools often seems to be the transfer of this substance" (Brown et al., 1989, p. 32). Brown et al. do not consider the declarative knowledge taught in schools to be robust, so that in continuing to focus on the transfer of such knowledge, the school culture becomes inauthentic, providing students with ersatz activities.

Schools should instead provide students with a context and activities that lead to the construction of knowledge as it used; the notion of performance assessment thus appears integral to this position.

Although there is no doubt that context-based learning is much more powerful than learning disembedded content, there is a place in school for learning decontextualized content; indeed, it could be argued that this is what schools are mainly for (Biggs, 1992a). If knowing and doing were inseparable, there could be a problem in accounting for civilization; to know only through doing virtually requires each generation to reinvent the wheel. Perhaps, then, we are talking about two different things: the status of different kinds of knowledge and the most efficient means of learning any kind of knowledge.

Whatever means of teaching we adopt—inductive, problem-based, hands-on on the one hand versus expository on the other—there is much knowledge, school-delivered, that is legitimately propositional or declarative and that needs to be assessed as such, as well as being assessed in its real world applications. Thus although learning takes place most easily in situated contexts, schools exist precisely to help students learn decontextualized second-order symbol systems, and the declarative and procedural knowledge encoded by them. Learning these systems and contents is unlikely to occur in children's direct experience of the world, and it does not come easily, possibly for biological reasons (Biggs & Moore, 1993), but such knowledge needs to be learned, and the quality of that learning needs to be assessed.

In short, then, there is a place for decontextualized learning and assessment, just as there is for situated. So far the assessment of decontextualized learning has been in practice greatly overemphasized quite out of proportion to its place in the regular curriculum, but it would be equally unbalanced to entertain only situated learning and assessment contexts.

Let me now refer to a model that provides a useful framework for both modes of assessment: for charting the course of developmental assessment and for providing a generalizable language for discussing student performance in at least some performance assessment tasks (I prefer my own term *ecological*, but *authentic* has only just been deposed and further semantic quibbling is undesirable at this stage).

A Generalized Model of Qualitative Assessment

In the developmental model of assessment it is first necessary to chart the course of development of a concept or principle, so that the stages of development can be defined and the level at which a student is currently thinking determined. We thus need to describe what the learning will be like at any particular stage in its growth. This may be done on a topic-by-topic basis, as has been for some topics by Marton and his co-workers using the techniques of phenomenography (Marton, 1988; Ramsden, 1988), which involves probing interviews that usually reveal layers of understanding of the target concepts: a hierarchy of conceptions that can be used to form assessment targets.

A more general model would require us to define increasingly higher quality in terms of such as aspects as increasing complexity of structure, abstractness, economy or elegance of processing, originality of the response, and so on. Quality involves many different aspects according to the task in question, but one aspect that is common to most tasks is the structural com-

plexity of the learning outcome. Basically, there are two aspects to structural complexity: the amount of detail in the student's response (the quantitative aspect), and how well put together that detail is (the qualitative aspect). Both aspects are important and may be classified by the SOLO Taxonomy (Biggs & Collis, 1982; 1989).

SOLO, which stands for *S*tructure of the *O*bserved *L*earning *O*utcome, provides a systematic way of describing how a learner's performance grows in complexity when mastering many tasks, particularly the sort of tasks undertaken in school. A general sequence in the growth of the structural complexity of many concepts and skills is postulated, and that sequence may be used to guide the formulation of specific targets or the assessment of specific outcomes.

1. The task is not attacked appropriately; the student hasn't really understood the point and uses too simple a way of going about it (prestructural).
2. One (unistructural), then several (multistructural), aspects of the task are picked up and used, but are treated independently and additively. Assessment of this level is primarily quantitative.
3. These aspects then become integrated into a coherent whole (relational); this level is what is normally meant by an adequate understanding of the topic. Assessment of this level becomes qualitative if it is to pick up its nature.
4. The previous integrated whole may be conceptualized at a higher level of abstraction and generalized to a new topic or area; this too requires qualitative assessment.

These levels and the general structures incorporated in each provide a basis for assessing the quality of particular learning episodes. It should be noted that although levels 1 and 2 may be assessed quantitatively and additively, 3 and 4 require an interpretive or hermeneutic approach (Moss, 1994); it is not the components themselves that determine the quality of the outcome, but the whole that their interaction and integration defines.

The SOLO taxonomy may be used in two kinds of assessment format:

1. *Assessing open-end outcomes* is a procedure that is straightforward and moderately well documented (Biggs & Collis, 1982). Its application to marking assignments in a standard letter-grade system is given in Biggs (1992b).

Although SOLO may not exactly provide a common currency across tasks, it does provide a common way of thinking about performance in quite different tasks, which is why it may be particularly useful in assessing, for example, students' portfolios. The justification of each performance the student provides in a portfolio, and the pattern formed by the selections as a whole, will exemplify a SOLO structure, which will say something about the way the student thinks about the course: as several unrelated tasks or performances, or as different key tasks that reveal an integrated way of conceptualizing the course, or even inventive applications or generalizations that go beyond the course itself.

2. *Assessing in an objective-type format*, the ordered-outcome format (Masters, 1987) is less well documented and needs some discussion here.

The ordered outcome format looks like a multiple-choice item without the choice: the subitems are ordered in a hierarchy of competence and all require a response, each indicating a particular level in the competence hierarchy.

Masters (1987) used a phenomenographic hierarchy in his prototype of the format (see above; Marton, 1988), but the disadvantage is that phenomenography is highly content specific and may yield any number of levels, which makes it complicated to use in practice. Using SOLO the following criteria emerge for subitems addressing the stem topic at each SOLO level:

1. *Unistructural*: Contains one obvious piece of information coming directly from the stem.
2. *Multistructural*: Requires using two or more discrete and separate pieces of information contained in the stem.
3. *Relational*: Uses two or more pieces of information each directly related to an integrated understanding of the information in the stem.
4. *Extended Abstract*: Requires use of an abstract general principle or hypothesis that can be derived from, or suggested by, the information in the stem.

To illustrate, an ordered-outcome mathematics test was given to several hundred Year 7 students in each of two Hong Kong schools (Biggs, Lam, Balla, & Ki, 1988). The content of the test is unremarkable in itself, but the responses to one item are revealing in the present context (see Figure 1).

The two schools perform similarly up to multistructural level, but they diverge sharply thereafter, 48% of School B students obtaining correct on the extended abstract subitem: 48% correct compared to 6% of School A; the differences between the students in Schools A and B, whatever their genesis might be, are reflected only in the most complex cognitive processes. A conventional test comprising an aggregate of mixed items scored correct or incorrect would be unlikely to pick up this qualitative difference in the students' mathematical thinking. Further, there is little reason why a teacher of grade 7 would normally think to test beyond a multistructural level. Here the ordered outcome format forces the teacher to think upward when designing the test, and hopefully likewise the student when responding.

Toothpicks are used to make the above patterns. Four are used to make one box, seven to make two boxes, etc.

		FORM 1	
		School A	School B
UNISTRUCTURAL	a. How many toothpicks are used to make three boxes ?	96%	99%
MULTISTRUCTURAL	b. How many more toothpicks are used to make 5 boxes than used to make 3 boxes ?	74%	76%
RELATIONAL	c. How many boxes can be made with 31 toothpicks ?	57%	70%
EXTENDED ABSTRACT	d. If I have made y boxes, how many toothpicks have I used ?	6%	48%

Figure 1. An ordered outcome mathematics item (Biggs et al., 1988).

In pointing out the two traditions in educational thinking, quantitative and qualitative, Cole (1990) points out that each has had a useful history and a continuing presence, but that as each appears incompatible the "public understanding of education is hurt by allowing these two unconnected conversations about educational achievement to continue separately" (p. 5), so that an integrating framework is necessary to allow each to persist where it is effective and appropriate. SOLO appears to provide such an integration: in the early learning of many topics, quantitative aggregation is an appropriate or convenient way of assessing, but for higher-order, applied, and critical thinking (relational and extended abstract) qualitative assessment is more appropriate. The important thing is that modes of assessment appropriate to lower levels do not preclude higher levels or suggest to students that they can meet requirements by substituting lower for higher levels, which is the usual problem with backwash.

Backwash: Problem or Solution?

I now return to the question of backwash: the notion that testing drives not only the curriculum, but teaching methods and students' approaches to learning, usually adversely (Crooks, 1988; Elton & Laurillard, 1979; Frederiksen, 1984; Frederiksen & Collins, 1989).

These observations on the effects of backwash have been largely of the traditional measurement model framework, or in CRT conceived quantitatively. In the last case, backwash is explicit and actively encouraged insofar as the test deliberately becomes the target for teaching, as in measurement-driven instruction (MDI) (Popham, 1987). An important value question is raised here, as there is evidence that the success of such a strategy depends on how students typically go about learning: those who typically focus on and reproduce detail (a surface approach to learning) like the strategy and do well, but those who adopt a more academic or deep approach, originally better than surface learners, become frustrated and do progressively worse (Lai & Biggs, 1994).

The problem here is simply that the target of learning defined in the quantitative framework is of a low cognitive level, covering only the first two levels in the SOLO taxonomy. Would targets of higher cognitive levels promote backwash beneficial to instruction and encourage students to adopt deeper approaches to learning? There is some evidence suggesting this to be so. Tang (1991), for example, showed that different modes of assessment, final (short-answer) examination and a single-topic assignment, elicited quite distinct assessment preparation strategies, the assignment generally producing higher cognitive level strategies based on wide reading, collaborative learning, and problem solving, but overriding the question of the actual mode of assessment was the student's perception of what was required for optimal results. Further, students needed to have the procedural knowledge necessary to enact those perceived requirements; extended writing was a novel task to many of Tang's students, and lack of the appropriate skills prevented them from realizing what they could see as required.

Higher cognitive-level targets need therefore to be perceived as requiring high cognitive-level preparation strategies that are in the student's repertoire. Ordered-outcome testing should send the message that what is important is to think in increasingly complex ways about a topic, not to obtain a certain

number of correct items. The gap between a student's best response and the highest response in the hierarchy tells both teacher and student what still remains to be learned. Wong (1994) set parallel forms of a grade 11 math test using the traditional format and quantitative scoring, and the ordered-outcome format; he then interviewed students while they solved the different item-types. The difference he describes is that between novices (solving correctly but algorithmically) and experts (solving economically and originally from first principles), yet it was the same student who was both novice and expert, the variable being the test item type. It is important that further such studies be carried out, as this would provide much needed empirical as opposed to rhetorical justification for qualitative approaches to assessment.

Dimensions of Assessment

Performance/authentic assessment, then, is but one aspect of the development from traditional testing to the current situation with respect to alternative modes of assessment. To put this in perspective, at least three dimensions are involved in this evolution:

1. The *function* of testing: is it to rank individuals along some assumed trait, as in Taylor's (1994) measurement model, or to refer an individual's performance to a standard?
2. The *nature* of what it is that is to be assessed: a unitized fragment of performance or the growth of understanding?
3. The *context* in which the test item is placed: is it embedded in a context isomorphic to that in which the knowledge is or will be used in everyday life, or is it abstract and decontextualized?

Table 1 puts these points together.

The qualitative-quantitative dimension establishes the nature of what is to be assessed and how it may be reported. The measurement and standards models may operate within the quantitative, yielding NRT and CRT as traditionally implemented, but only the standards model may operate within the qualitative dimension, because NRT requires unidimensionality, whereas assessing an outcome qualitatively is a hermeneutic not a dimensional or measurement-based process (Moss, 1994). The original authentic debate raised the issue of whether assessment might best be situated or decontextualized. Decontextualized assessment would include the usual method of pencil-and-paper testing, which may be construed within either a quantitative framework, as is traditionally has been the case, or a qualitative framework, as is the case with most SOLO testing to date and with ordered-outcome testing. We then need to distinguish between testing the student's developing understanding of

Table 1
Dimensions and Modes of Assessment

	Model	*Context (examples)*	
		Decontextualized	*Situated*
Quantitative	Measurement	1. NRT (MC test)	4. PA (NRT) (?)
	Standards	2. CRT (Mstry Lrn)	5. PA (CRT)
Qualitative	Standards	3. SOLO (Develptl)	6. PA (Ecolgl)

a concept, particularly but not essentially of declarative knowledge, and the student's ability to involve that knowledge in a task that has ecological validity with respect to the learning goals. The issue in the last case is not so much what kind of understanding students have of the content, but whether the taught content can empower decision making in a real context (Maguire, 1990; Masters & Hill, 1988; Wiggins, 1989).

We are, then, left with six cells indicating modes and contexts of assessment.

1. *Quantitative-Measurement-Decontextualized*: Traditional NRT, which now has a sophisticated technology and the overwhelming approval of the measurement establishment and of most administrators. Best used for selection and other individual or group comparative purposes, relatively curriculum-free. The backwash generated is almost certainly the most deleterious to teaching and learning of all six modes.
2. *Quantitative-Standards-Decontextualized*: 1970s style CRT, mastery learning. Linked closely to curriculum and instructional strategy. Best used for testing basic or core skills, behavioral skills. The quantitative framework, however, limits its application to lower cognitive levels. Backwash suits students preferring a surface approach to learning, but counterproductive for students preferring a deep approach (Lai & Biggs, 1994). A model that in many senses straddles 1 and 2 is Item Response Theory (IRT) or Rasch Model (Wright & Stone, 1979), which is quantitatively framed and uses a version of trait theory, but the scoring does not depend on how other students perform as in NRT, or on the items making up the test as in CRT. However IRT is controversial even within the quantitative framework, but is open to most of the same criticisms.
3. *Qualitative-Standards-Decontextualized*: The developmental mode of assessment where the interest is in the growth of skills and concepts in themselves rather than in their applications, using pencil-and-paper testing rather than situated performances. Best used for finding the levels of understanding of basic concepts that students have attained so far; these levels could (and should) be incorporated into the curriculum as targets for instruction. Alternative framework research (White, 1988) and phenomenography (Marton, 1988) use situations and hierarchies specific to given topics, whereas SOLO uses a more general framework. Ordered-outcome testing, whether based on SOLO or any other growth model, also falls into this category. Backwash from such testing is likely to be helpful as it encourages teacher and learner to think higher.
4. *Quantitative-Measurement-Situated*: PA in NRT mode. This would involve setting tasks the grading of which is competitive; perhaps an athletics meet is an example. However, the assumption of the measurement model, that individual scores can be related to a stable trait, is really not what PA is about. For all intents and purposes, then, this is an empty cell.
5. *Quantitative-Standards-Situated*: many PA tasks could be scored and graded quantitatively, for example, number of correct applications or behaviors, use of rating scales. Such quantitative assessments are useful for communication or for manipulation in combining grades or determining cut-offs. Although this may seem open to some of the objections already raised against quantitative assessment, it is much clearer to both teachers and

learners that the numbers are used here for logistic purposes and are not an implied statement about the quantitative nature of the performance. The form of understanding being created by the context is appropriate to the real-world use of the knowledge in question precisely because it is situated.

6. *Qualitative-Standards-Situated*: Ecological PA where the assessment task is situated and evaluated in a context close to the learning goals. Grading would be in terms of qualitative categories of mastery or competence. These might be quite task specific, or based on a more general framework such as SOLO. Best used for learnings that are meant to have direct real-world applications, which accounts for much school learning but not all. Backwash is likely to be helpful for learning as the testing situation should be designed to mimic the learning objectives, thus excluding cynicism and minimizing test-wiseness, short-cuts, or surface learning.

No doubt other modes of assessment could be generated using different or additional dimensions. Perhaps one that is missing here is self- versus other-assessment; that is a particularly interesting pivot in the *situated* column. Nevertheless, the six, or rather five, cells generated here show a disproportion in terms of frequency of use in relation to the instructional context in which they are most useful.

Part of this disproportion is due a context that is inimical to the newer, alternative modes of assessment. The measurement establishment has in effect imposed a set of assumptions about the nature of educational measurement that is simply at odds with what should go on in classrooms; administrators have, moreover, encouraged this domination because the assumptions and practices of traditional measurement serve political and utilitarian administrative ends rather than educational ones (Wilson, 1994). Nowhere could this be clearer than in the current pressures toward competence-based testing currently being experienced in many countries (Biggs, 1994). The quantitative-decontextualized framework of assessment, whether CRT or NRT, has simply been hard to resist, either because of direct pressure or more simply from inertia.

Nevertheless, as the rapidly increasing interest and acceptance of PA shows, it is clear that the time has come for change, and increasingly teachers are recognizing this, perhaps more so in Canada (Bachor & Anderson, 1994) than in other countries. A real practical difficulty in the way of wider acceptance is not that practitioners see no need for change, but that they lack a framework for interpreting results and incorporating them into established grading schemes; the reliability and validity of the new techniques are not evident, although many forms of assessment and in particular report writing provide a heavy workload (Bachor & Anderson, 1994). The cost benefits of PA are simply not seen to be favorable. All this suggests that much further work is necessary.

What needs to be done to make PA and qualitative assessment in general more acceptable?

Some Current Problems

Two main issues warrant discussion: determining conceptions of reliability and validity acceptable to practitioners, and adapting institutional structures to accept and make performance assessment workable in classroom contexts.

Conceptions of Validity and Reliability

As noted, the assumptions underlying PA and qualitative assessment in general are quite different from those underlying the measurement model. It is therefore inappropriate to apply the same tests and standards of reliability and validity. Messick's classic (1989) review of validity set the tone for rethinking this issue, but this article is not particularly oriented to qualitative assessment. However, his central point, that validity is not a property of the test, but of the interpretations and uses to which test scores are put and their consequences, opened the way to considerable discussion and a good deal of consensus about reliability and validity specifically in relation to PA (Bachor, Anderson, Walsh, & Muir, 1994; Frederiksen & Collins, 1989; Haertel, 1991; Linn, Baker, & Dunbar, 1991; Messick, 1994; Moss, 1992, 1994; Shepard, 1993; Taylor, 1994; Wolf, Bixby, Glenn, & Gardner, 1991).

Perhaps the most fundamental insight is that validity is now seen as being grounded in the theory and knowledge base of learning and teaching, not in a science and technology of educational measurement, so that "in place of ranks, we will want to establish a developmentally ordered series of accomplishments" (Wolf et al., 1991, p. 63). The task of test construction, then, becomes rather different from the traditional model; one now requires a theory of learning unfolding longitudinally, the construct being tested well represented in the elicited test behaviors and explaining them and specific stepped targets for each curriculum topic. The question is no longer, Does a test measure what it is supposed to measure? but, Does it do what it is supposed to do? (Shepard, 1993). Thus assessment becomes contextualized, so that although accuracy, coverage, or representativeness of teaching goals in the test items and consistency are important, so too are fairness and adverse consequences, although whether these last questions belong in the domain of test validity or professional ethics in general is debatable, although their importance is not (Maguire, Hattie, & Haig, 1994).

Another fundamental shift from classic test theory is the role played by judgment and consensus both in establishing construct validity (an end to those days when a test was "valid for anything with which it correlates" [Shepard, 1993]), and in interpreting test scores. Thus, although many are concerned about the low reliability of portfolios, for example, Bateson (1994), it is important to recognize that the old additive model of compounding unreliability no longer applies. Specifically, a hermeneutic approach to drawing conclusions from test performances avoids this problem, as the aim is to arrive at a judgment by understanding the whole in light of the parts; it is not a case of judging single performances and then aggregating. An example is how a journal editor judges whether to accept or reject a manuscript on the basis of informed advice, even when rejecting a paper advocating a hermeneutic over an additive approach on the grounds that additivity is the accepted model (Moss, 1994).

Finally, both fidelity or ecological validity and fairness require that multiple routes to the same goal performance are allowed. As in a music competition, performers choose their own particular items, often using different instruments in order to show themselves at their best; analogically, then, students have the right to choose what to put in their portfolio (Moss, 1994). The final

decisions, whether summative or formative, are holistic and qualitative, requiring expert judgment.

For educators brought up on traditional test theory these considerations are likely to feel alien if not misleading. Undoubtedly Bateson (1994) is correct in saying "University preservice training has generally failed dismally in preparing teachers for the testing, assessment, and evaluation tasks which they must undertake in their classrooms" (p. 239). However, there is now a growing theory, and its technology, that could inform preservice and inservice teacher education and could replace the unsystematic but "amazing 'moccasin telegraph'" (p. 240) that spreads new and good ideas among teachers.

Institutional and Systemic Issues

The commitment and competence of teachers is certainly important in impeding change, but equally if not more important is the way institutions run. An educational institution is a system, that is, a working whole made up of a set of component parts, each of which affects the other until the whole forms an equilibrium, that state of equilibrium becoming the system (von Bertalanffy, 1968). The prime example of a system is, of course, an ecological system, in which individual component organisms lie in a delicate state of symbiotic interdependence. However, systems theory can apply to almost any complex situation, including schools and classrooms (Yinger & Hendricks-Lee, 1993). In the classroom the system comprises teacher, students, curriculum, teaching method, methods of assessment, and learning and teaching outcomes, all in a state of interdependence (Biggs, 1993). Part of the fierce resistance to change in assessment schemes is that assessment is so powerful a component, its backwash affecting the preceding chain of teaching and learning. Unless all participants are willing to change in the directions required, resistance will be considerable. The British Columbian situation (Bachor & Anderson, 1994; Bateson, 1994) illustrates both the difficulty and the possibility of change.

Reid (1987) points to three important components in the institutional system itself: the rhetoric, or the official aims of teaching; the technology, which would make possible the realization of these aims; and the social system of the institutions, which determines what is allowable in the institution. The social system comprises: the requirements established on a collegial basis, mostly informal but often given formal weight in Faculty meetings; the formal requirements of bureaucracy; and the requirements of the student body, which may be formal or informal. It is probably the social system with its various collegial, accountability, and managerial agendas that exerts most pressure on the assessment system in use (Biggs, in press).

The greatest pressures will be for using quantitative assessment. Most institutions require combining summative assessments at some stage: subunit results to obtain course results, course results to obtain year results. This puts almost irresistible pressure on teachers to use quantitative marking schemes, because marks are easily added up and averaged and make discrimination between students extremely easy. Profiling or other qualitative schemes could be used, but usually are not in the event. Most teachers, therefore, mark quantitatively, with the sort of results discussed above (Lohman, 1993).

Pressures toward decontextualized testing also exist in terms of convenience, security, and tradition. Institutionally, administrators feel it essential

to have standardized procedures, timed testing conditions, and cheat-proof security in the interests of fairness and accreditation and in anticipation of possible law suits. The public, and students too, see a face validity in this. As for teachers, it is difficult enough breaking the habits of a lifetime to redesign situated assessment, let alone under the conditions prescribed by the bureaucracy.

Thus changing the assessment system means setting up a new equilibrium, perhaps requiring a new technology, almost certainly requiring a new deal to be struck at all levels in the existing social system of the institution: with colleagues, with the bureaucracy, and with students. It is not impossible, as the British Columbian experience shows, but it is likely to be difficult.

Conclusions

It is no exaggeration to say that the theory and practice of assessing learning are currently undergoing a major paradigmatic change. It is not a matter of CRT replacing NRT, but the intersection of a variety of movements that, interestingly, have come from the broader educational canvas, not from the testing establishment itself.

Certainly the notion of CRT is coming to the forefront, but in connection with other views about the qualitative nature of higher order learnings and the situated nature of learning. Few of these ideas are particularly new, but their interaction becomes paradigmatic, suggesting heavily revised views of reliability and validity and new formats of testing. The critical single notion underlying this is, amazingly enough, that educational considerations should drive testing, and not psychometric, bureaucratic, or political ones.

Assessment occupies a key place in determining quality learning outcomes, but assessment practices are part of a wider picture that includes but extends beyond the responsibility of any individual teacher. An institution is a holistic, interactive system, which for its own management has many procedures in place with their own functional use. However, these procedures often determine teaching and assessment practices, which in turn influence students' perceptions of what and how they will learn.

There are three main factors impeding change. The first and simplest is know-how; many teachers may simply not know how to improve their assessment techniques. Second, and more subtle, is the probability that they don't know that they don't know. If someone has a quantitative mindset, really believing that we should teach, learn, and assess by numbers, then they won't even see that there is a problem. Finally, teachers have the institutional social system to deal with.

Obviously the road to better teaching, learning, and assessment is a complicated one that is beyond the control of educational researchers themselves. However, the history of assessment has shown the unfortunately negative effect the outdated measurement establishment has had on classroom practice. At least we can do something about that, and systems being what they are, perhaps we can then strike different and healthier equilibria than currently exist.

Acknowledgment

I am indebted to John Kirby, Bob Wilson, and one anonymous reviewer for their valuable feedback and advice.

Note

This article is based on an invited address to the Canadian Association for Educational Psychology, presented at the CSSE conference, Calgary, June, 1994.

References

Archbald, D.A., & Newman, F.M. (1988). *Beyond standardized testing: Assessing authentic achievement in the secondary school*. Reston, VA: National Association of Secondary Principals.

Bachor, D.G., & Anderson, J.O. (1994). Elementary teachers' assessment practices as observed in the province of British Columbia. *Assessment in Education, 1*, 63-93.

Bachor, D.G., Anderson, J.O., Walsh, J., & Muir, W. (1994). Classroom assessment and the relationship to representativeness, accuracy, and consistency. *Alberta Journal of Educational Research, 40*, 247-262.

Bateson, D.J. (1994). Psychometric and philosophic problems in "authentic" assessment: performance tasks and portfolios. *Alberta Journal of Educational Research, 40*, 233-246.

Biggs, J.B. (1973). Study behaviour and performance in objective and essay formats. *Australian Journal of Education, 17*, 157-167.

Biggs, J.B. (1992a). Returning to school: Review and discussion. In A. Demetriou, M. Shayer, & A. Efklides (Eds.), *Neo-Piagetian theories of cognitive development* (pp. 277-294). London: Routledge and Kegan Paul.

Biggs, J.B. (1992b). A qualitative approach to grading students. *HERDSA News, 14*(3), 3-6.

Biggs, J.B. (1993). From theory to practice: A cognitive systems approach. *Higher Education Research and Development, 12*, 73-86.

Biggs, J.B. (1994). Learning outcomes: Competence or expertise? *Australian and New Zealand Journal of Vocational and Educational Research, 2*, 1-18.

Biggs, J.B. (in press). Assessing learning quality: Reconciling institutional, staff and educational demands. *Assessment and Evaluation in Higher Education*.

Biggs, J.B., & Collis, K.F. (1982). *Evaluating the quality of learning: The SOLO Taxonomy*. New York: Academic Press.

Biggs, J.B., & Collis, K.F. (1989). Towards a model of school-based curriculum development and assessment: Using the SOLO Taxonomy. *Australian Journal of Education, 33*, 149-161.

Biggs, J.B., Lam, R.Y.L., Balla, J.R., & Ki, W.W. (1988, November 26-27). *Assessing learning over the long term: The "Ordered Outcomes" model*. Paper presented to Annual Conference, Hong Kong Educational Research Association.

Biggs, J.B., & Moore, P.J. (1993). *The process of learning*. Sydney: Prentice-Hall Australia.

Bloom, B.S., Hastings, J.T., & Madaus, G.F. (1971). *Handbook of formative and summative education of student learning*. New York: McGraw-Hill.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-41.

Cizek, G. (1993). Rethinking psychometricians' beliefs about learning. *Educational Researcher, 22*(4), 4-9.

Clark, J., Scarino, A., & Brownell, J. (1994). *Improving the quality of learning: A framework for target oriented curriculum renewal in Hong Kong*. Hong Kong: Honkong Bank Language Development Fund, Institute of Language in Education.

Cole, N.S. (1990). Conceptions of educational achievement. *Educational Researcher, 19*(3), 2-7.

Crooks, T.J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*, 438-481.

Delandsheere, G., & Petrovsky, A.R. (1994). Capturing teachers' knowledge: Performance assessment. *Educational Researcher, 23*(5), 11-18.

Elton, L., & Laurillard, D. (1979). Trends in student learning. *Studies in Higher Education, 4*, 87-102.

Frederiksen, J.R. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist, 39*, 193-202.

Frederiksen, J.R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher, 18*(9), 27-32.

Haertel, E.H. (1991). New forms of teacher assessment. *Review of Research in Education, 17*, 3-29.

Hoffman, B. (1962). *The tyranny of testing*. New York: Collier.

Lai, P., & Biggs, J.B. (1994). Who benefits from mastery learning? *Contemporary Educational Psychology, 19, 13-23.*

Linn, R., Baker, C., & Dunbar, S. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15-21.

Lohman, D.F. (1993). Teaching and testing to develop fluid abilities. *Educational Researcher, 22*(7), 12-23.

Maguire, T.O. (1990, May 31-June 1). *Grounded authentic assessment and teacher education.* Paper presented to the Second Conference on Classroom Assessment, Vancouver.

Maguire, T.O., Hattie, J., & Haig, B. (1994). Construct validity and achievement assessment. *Alberta Journal of Educational Research, 40,* 109-126.

Marso, R.N., & Pigge, F.L. (1991). An analysis of teacher-made tests: Item-types, cognitive demands, and item construction errors. *Contemporary Educational Psychology, 16,* 279-286.

Marton, F. (1988). Describing and improving learning. In R.R. Schmeck (Ed.), *Learning strategies and learning styles.* New York: Plenum.

Marton, F., Dall'alba, G., & Beaty, E. (in press). Conceptions of learning. *International Journal of Educational Research.*

Masters, G. (1987). New views of student learning: Implications for educational measurement. *Research working paper 87.11.* University of Melbourne, Centre for the Study of Higher Education.

Masters, G.N., & Hill, P.W. (1988). Reforming the assessment of student achievement in the senior secondary school. *Australian Journal of Education, 32,* 274-286.

Messick, S.J. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18*(2), 5-11.

Messick, S.J. (1994). The interlay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13-23.

Moss, P.A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Review of Educational Research, 62,* 229-258.

Moss, P.A. (1994). Can there be validity without reliability? *Educational Researcher, 23*(2), 5-12.

Newman, F.M., & Archbald, D.A. (1992). The nature of authentic academic achievement. In A.R. Tom (Ed.), *Toward a new science of educational testing and achievement.* Albany, NY: State University of New York Press.

Popham, W.J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappan, 68,* 679-682.

Popham, W.J., & Husek, T.R. (1969). Implications of criterion-referenced measurement. *Journal of Educational Measurement, 6,* 1-9.

Ramsden, P. (Ed.). (1988). *Improving learning: New perspectives.* London: Kogan Page.

Reid, W.A. (1987). Institutions and practices: Professional education reports and the language of reform. *Educational Researcher, 16*(8), 10-15.

Shepard, L.A. (1991). Psychometricians' beliefs about learning. *Educational Researcher, 20*(6), 2-16.

Shepard, L.A. (1993). Evaluating test validity. *Review of Research in Education, 19,* 405-450.

Tang, K.C.C. (1991). *Effects of different assessment methods on tertiary students' approaches to studying.* Unpublished doctoral dissertation, University of Hong Kong.

Taylor, C. (1994). Assessment for measurement or standards: The peril and promise of large scale assessment reform. *American Educational Research Journal, 31,* 231-262.

Trigwell, K., Prosser, M., & Taylor, P. (1994). Qualitative differences in approaches to teaching first year university science. *Higher Education, 27,* 75-84.

Von Bertalanffy, H. (1968). *General systems theory.* New York: Braziller.

White, R.T. (1988). *Learning science.* Oxford, UK: Blackwell.

Wiggins, G. (1989). Teaching to the (authentic) test. *Educational Leadership, 46,* 41-47.

Wilson, R.J. (1994, June). *Back to basics: A revisionist model of classroom-based assessment.* Presidential address, Annual Meeting, Canadian Educational Researchers Association, CSSE, Calgary.

Wilson, R.J., & Kirby, J.R. (1994). Introduction: Special issue on cognition and assessment. *Alberta Journal of Educational Research, 40,* 105-108.

Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. *Review of research in Education, 17,* 31-74.

Wong, C.S. (1994). *Using a cognitive approach to assess achievement in secondary school mathematics.* Unpublished master's thesis, University of Hong Kong.

Wright, B.D., & Stone, M.H. (1979). *Best test design: Rasch measurement.* Chicago, IL: MESA Press.

Yinger, R.J. & Hendricks-Lee, M.S. (1993). An ecological conception of teaching. *Learning and Individual Differences, 5,* 269-282.

The Alberta Journal of Educational Research *Vol. XLI, No. 1, March 1995, 18-35*

Debora Barnett-Foster
Ryerson Polytechnic University

and

Philip Nagy
Ontario Institute for Studies in Education

A Comparison of Undergraduate Test Response Strategies for Multiple-choice and Constructed-response Questions

In this study the response strategies employed by a group of undergraduate chemistry students to answer multiple-choice and stem-equivalent constructed-response test questions were examined. A combination of written (N=300) and oral (N=21) reports from the students were analyzed. The written data were analyzed for format-related differences in solution strategies and error patterns. The oral reports were scored for the presence of typical problem solving behaviors. Analysis of the written data indicated no significant differences in the types of solution strategies employed nor in the types of errors committed across test format. The oral data, however, revealed evidence of subtler differences in the types and frequencies of problem solving behaviors.

Cette étude examine les différentes stratégies utilisées par un groupe d'étudiant(e)s en chimie au niveau universitaire pour répondre à des questions à choix multiples et à des questions construites selon un modèle où il y a plusieurs réponses possibles à une question-souche. On analysa une combinaison des réponses écrites (N=300) et des réponses orales (N=21) obtenues des rapports des élèves. Les données écrites ont été analysées pour relever les différences reliées au format des stratégies des solutions et des tendances des erreurs trouvées. Les rapports oraux ont été évalués pour la présence des comportements typiques concernant la résolution de problèmes. L'analyse des données écrites indique qu'il n'y a aucune différence significative dans les sortes de stratégies utilisées ni dans les sortes d'erreurs produites par les élèves entre les deux différents genres de tests. Les données orales cependant ont révélé la présence de différences plus subtiles dans les sortes et les fréquences de comportements utilisés en résolution de problèmes.

Introduction

The question of test format equivalence has been and continues to be an issue of some controversy. Since the evolution of the multiple-choice test, educational researchers have debated the relative merits of this format versus formats that require the student to compose or construct an answer. Comparisons of test format have involved studies of reliability, validity, success rate, item difficulty, knowledge retention effects, expectancy effects, and strategies for preparing for tests. However, one of the primary issues of interest has been the

---

Debora Barnett-Foster is a professor of applied chemical and biological sciences. Her interests are evaluation in undergraduate science education and test format equivalence.
Philip Nagy is an associate professor of measurement and evaluation. His interests are assessment of complex learning outcomes and data analysis.

investigation of the equivalence of traits or abilities measured by the tests. Although this has been an active area of research, the findings have been inconclusive. Some reports suggest that the formats measure different traits (Ackermann & Smith, 1988; Birenbaum & Tatsuoka, 1987; Ward, Frederickson, & Carlson, 1980). Yet others have found no effect of test format on the traits measured (Bennett et al., 1990; Bennett, Rock, & Wang, 1991; van den Bergh, 1990; Ward, 1982).

Traub (1992) argues that format effects can be expected in situations where the answers to test items provided in multiple-choice questions can be "recognized" or arrived at by some process of comparison or elimination not available to examinees answering constructed-response items. On the other hand, he suggests that items that require manipulation of data or ideas are not subject to format effects, because presumably examinees would have to perform these manipulations regardless of item format. If this is true, then the knowledge domain and the task demand under evaluation may influence the outcome of trait-equivalence research on the two test formats and may help to explain some of the apparent contradictions in the literature.

In most studies of trait equivalence the effect of format has been determined through statistical analysis of performance scores based on written test answers. Relatively little work has been done to identify and characterize the traits measured by the two test formats. Where it has been undertaken it has generally been confined to one type of test format and has often been based solely on written test answers. In most studies that have identified test response strategies employed to answer multiple-choice questions, the primary source of empirical data has been derived from statistical analysis of keyed options and distracters. There have been only a few investigations in which the reasoning processes were examined using think-aloud interviews (Bloom & Broder, 1950; Connolly & Wantman, 1964; Farr, Pritchard, & Smitten, 1990; Kropp, 1956; Norris, 1992). In the case of the constructed-response test format, most conclusions about test response strategies have been derived from statistical analysis of test subscores and error patterns in written test answers.

The commonly held assumption is that the constructed-response format requires the student to use some form of production strategy whereas the multiple-choice format requires only that the student discriminate among alternatives. However, the empirical evidence to date about these types of strategies is limited. The elimination or compare-and-delete strategy is often considered to be a predominant response strategy in answering multiple-choice questions; yet it has rarely been documented (Farr et al., 1990; Snow, 1980). It has also been suggested that the provision of answer options in the multiple-choice question might alter or truncate the solution path that would otherwise be used if there were no answer options to guide the solver. The solver may even work backward from the answer options to develop a solution path in a manner similar to that described by Hayes (1981). Although there has been considerable speculation about the type of response strategies employed to answer multiple-choice and constructed-response questions, there is very little empirical evidence to either support or refute these beliefs.

This investigation was intended to explore and describe the kinds of response strategies that students typically employ in answering questions posed

in multiple-choice and stem-equivalent constructed-response formats, using written and think-aloud procedures designed to elicit evidence of the students' thinking processes. In order to control for the potentially confounding effects of domain and task demand variables, this investigation was restricted to a single type of task in a specific discipline. Chemistry was selected because little work on format effects has been conducted in the sciences and none in the field of chemistry. In this discipline, tasks that involved quantitative problem solving were selected as being most typical of questions posed to first-year chemistry students.

Method

Overview

A combination of written test answers and oral reports of the examinees about their answers were used to compare the types of response strategies employed across test format. The written test answers were categorized and analyzed for differences in solution strategies and error patterns across test format. The oral reports were coded and analyzed for the thought processes, particularly with regard to problem solving behaviors associated with written solution strategies.

Subjects

A class of 300 students in a first-year introductory chemistry course was selected for this study. All students in the course had at least one secondary school credit in chemistry. Students in the course were registered in one of three undergraduate professional degree programs: Applied Chemistry and Biology, Environmental Health, or Food and Nutrition. The investigators neither taught the course nor had any other instructional role in it.

Three weeks prior to the final examination the students were informed of the purpose and nature of the project and were invited to participate; 261 agreed to allow their test answers to be analyzed. From that group 51 students volunteered to participate in a posttest interview. Of that group 21 were selected to represent three ranges of academic achievement based on secondary school chemistry grades: 76-99% (n=6), 65-75% (n=9) and below 65% (n=3). In addition three students who had been admitted to the institution as mature students were also interviewed.

Instrument

The test instrument was the final examination of the course. As shown in Table 1, the test was prepared in two versions that were distributed to randomly divided halves of the class.

Each version consisted of 15 multiple-choice questions and five constructed-response questions. Some of the multiple-choice questions (designated MCQ) simply requested the examinee to answer by selecting one of the five multiple-choice options provided. Other multiple-choice questions (designated MCQWR) requested the examinee to select one of the five options and to provide an explanation for the selected answer. The constructed-response questions (designated CRQ) asked the examinee to compose a response. For reasons of examinee convenience, and in order to separate the parallel questions as much as possible, the actual order of item presentation does not correspond to the order of the numbered sets. In all cases the multiple-choice

Table 1
Test Construction Design

Question Set	Content Area	Cognitive Level	Question Types Test A	Test B
1 (4 items)	molarity	(RCA)	MCQ	MCQ
	buffers	(CCA)		
	stoichiometry	(CCA)		
	kinetics	(RCA)		
2 (3 items)	molarity	(RAA)	MCQ	MCQWR
	kinetics	(RRC)		
	thermodynamics	(RRC)		
3 (3 items)	half-life	(RCA)	MCQWR	MCQ
	thermodynamics	(CCC)		
	thermodynamics	(AAA)		
4 (5 items)	gas laws	(CAA)	MCQWR	CRQ
	kinetics	(AAA)		
	crystal structure	(AAA)		
	stoichiometry	(CAA)		
	buffers	(AAA)		
5 (5 items)	gas laws	(CAA)	CRQ	MCQWR
	kinetics	(AAA)		
	crystal structure	(AAA)		
	stoichiometry	(CAA)		
	buffers	(AAA)		

Note 1. MCQ, MCQWR and CRQ represent multiple-choice questions, multiple-choice questions with response and constructed-response questions respectively. In a single row, the question sets are identical except for format.
Note 2. The cognitive level assigned by the three judges for each question is provided in column 3 as a series of three letters: R for Fact Recall, C for Comprehension, and A for Application.

items (MCQ and MCQWR) were presented before the constructed-response items.

The first question set (set 1) consisted of four multiple-choice questions (MCQ) that were identical across the two test versions. Question set 1 was designed to reveal performance differences in the two groups.

Question sets 2 and 3 were intended to provide an opportunity to examine the effect of requesting written explanations for the multiple-choice answers on the answer selection process. In set 2 the students who wrote test version A answered three questions in an MCQ format whereas the students who wrote test version B answered the same three questions in an MCQWR format. This pattern was reversed for question set 3, using a new set of three questions.

Question sets 4 and 5 were designed to provide a comparison of response strategies of students to the same questions posed in multiple-choice (MCQWR) and stem-equivalent constructed-response (CRQ) format.

Design of the Test Questions

Each multiple-choice question consisted of one correct response and four incorrect distracters. The distracters were chosen from a set of the most common errors produced by students tested with an equivalent constructed-response question. The errors were derived from archived examination papers and from a pilot study. Stem and option statements were examined to eliminate grammatical inconsistencies, inequality of length, and word cuing. Options that may have generated ambiguity such as "all of the above" or "A and B" were not used. Placement of correct responses was distributed over the five option positions so as to minimize serial position placement effects. Numerical answer options were sequenced in order of increasing value. An equivalent set of short-answer questions was developed using identical question stems.

Because the cognitive demand of a test question may affect test response behavior, the questions in this study were categorized according to the cognitive demand required to answer them. Questions were classified as requiring *fact recall, comprehension, or application* (Bloom, 1956) independently by three chemistry professors.

Questions belonging to sets 1, 2, and 3 were classified as fact recall (R), comprehension (C), or application (A) questions. All the questions from sets 4 and 5 were classified as application questions.

Administration of the Final Examination

Test administration conditions and instructions were identical for both groups and were consistent with typical examination procedures. Both groups were tested simultaneously. For MCQWR questions students were instructed to select an answer from the five options and to explain in writing their reasons for their answer selection. Students were also informed that the provision of an explanation was required in order to receive credit for the correct response. Students were permitted to answer the questions in any sequence they wished and were informed that there was no penalty for guessing.

Interview Procedure

After the examination 21 students attended a think-aloud interview with one of the investigators. The majority of interviews were conducted in a seven-day period following the written test. However, conflicts occasioned by examination and holiday schedules forced four interviews to be conducted one month after the written test.

At the beginning of each interview the think-aloud method was explained to the student, using Hayes, Flower, Schriver, Stratman, and Carey (1987) as a basis. Following this explanation a warm-up exercise was used to acquaint the student with the think-aloud technique. After the warm-up exercise the student was asked to read a handout that described the task about to be undertaken.

The student was then provided with an unmarked copy of his or her answers to the final exam, along with examination materials (periodic table, data sheet, calculator). The student was directed to examine certain test questions and to report, without attempting to explain or edit, all thoughts that had passed through his or her mind during the original problem solving session. In the event that the student could not recall how the problem had been solved

during the examination period, he or she was then instructed to attempt the problem and to vocalize all thoughts during this exercise. The provision of the written test responses during the oral protocols minimized the incidence of failure to recall.

During the interview the investigator intervened as little as possible. If the student paused for 10 seconds or longer, the investigator encouraged the student to continue to vocalize all thoughts. Otherwise the interviewer remained silent.

The meeting was tape-recorded and the investigator took notes. The sessions typically lasted one hour. At the end of the meeting all notes made by both parties were collected and collated with the copy of the subject's test.

The selection of test questions for interviews was designed to include for each subject three questions from each of set 2/3 and three from set 4 and set 5. Question selection was systematically varied across interviews so as to collect data on all questions in sets 2, 3, 4, and 5 for both test versions.

Results and Analysis

Analysis of Written Tests

For the purposes of this investigation, test responses for MCQ, MCQWR and CRQ were scored dichotomously. Differences in mean score and response pattern were examined for all written responses. Responses to question sets 4 and 5 were also analyzed for similarities and differences in the solution process used to generate the answer. The responses were first sorted on the basis of whether sufficient information was provided to categorize the solution processes used. Then those responses that were sufficiently complete were analyzed by question-content area for patterns in solution strategies and errors. The patterns were contrasted across question format.

Reliability coefficients (KR 20) were similar for both test formats. For the 15 multiple-choice items, the KR 20 values with means and standard deviations for versions A and B were 0.70 (8.6, 2.9) and 0.65 (8.7, 2.8) respectively. For the five constructed-response questions, the KR 20 values (means and standard deviations) for versions A and B were 0.74 (2.6, 1.7) and 0.64 (2.9, 1.5) respectively. It should be noted that the results for a single topic and format were combined across both test versions for analysis.

Given the sequential nature of some of the data analysis, with each subsequent stage being dependent somewhat on the results of the preceding stage, the description of the analysis and the presentation and discussion of the results are presented for each stage in order. Results of the written test analysis are reported by question set.

Question set 1. A student's t-test revealed a significant difference between the mean total scores computed from the four questions of set 1 for the two groups ($\bar{X}_{1=}$ 2.1, $\bar{X}_{2=}$ 2.4, $p \leq 0.01$). Using the two sample binomial test of equal proportions, analysis of performance differences between the two groups on each question revealed a significant difference only in the case of the buffer solution question. Because the mean score difference was only 0.3 of the pooled standard deviation (0.94) and a significant difference was found on only one question, the scores for sets 2 to 5 were not adjusted.

Question sets 2 and 3. A t-test analysis of the mean total scores for question set 2 (MCQ for test A and MCQWR for test B, identical questions) revealed no

Table 2
Response Choice Data (percent) for MCQ versus MCQWR (question set 3)

Question	Answer Option A	B	C	D	E	N/A
Half-life						
MCQ	4	79*	5	7	4	1
MCQWR	7	77*	5	6	3	2
Thermodynamics						
MCQ	0	1	56*	18	24	1
MCQWR	6	2	55*	17	21	0
Thermodynamics						
MCQ	3	39	39	10*	5	4
MCQWR	3	45	29	8*	4	11

*correct response.
N/A: no answer provided.

significant difference in performance between the two groups ($\overline{X}_1$= 1.4, $\overline{X}_2$= 1.5, p=0.15). A similar analysis for question set 3 also showed no significant difference in the mean total scores between the two groups (X_1=2.0, X_2=2.0). Because analysis of question sets 2 and 3 produced remarkably similar results, only the percent responses for set 3 are reported here in Table 2.

It is interesting to note that although there appeared to be a significant difference in the performance scores of the two groups for question set 1, the scores for sets 2 and 3 were not significantly different across the two groups. In the light of the crossed-design nature of sets 2 and 3, this finding provided support for the decision not to adjust scores for sets 2 to 5 based on the difference determined for set 1.

Chi-square analysis was used to determine differences in the frequencies of incorrect options chosen in question sets 2 and 3. The intent was to examine the effect of requesting a student to provide a written justification or reason for selecting his or her response. In particular, did this format lead to a change in the distribution of incorrect options selected? Chi-square values were determined for each question, first, with the four incorrect options and the no-answer response and, second, with the four incorrect options only. Because the two sets of data (including and excluding the no-answer response) differed only marginally, we decided to report only the results for the patterns of incorrect options selected. The Chi-square values revealed no significant differences in the patterns of incorrect responses across format for the six questions in sets 2 and 3 (Glass & Hopkins, 1984).

Question sets 4 and 5. Question sets 4 and 5 contained equivalent sets of multiple-choice and constructed-response questions designed to reveal differences in the response patterns for multiple-choice versus constructed-response questions. A two way (test x format) analysis of variance revealed no significant interaction between test and format (p=0.141). Test means for versions A and B were 2.79 and 2.89 respectively and format means for MCQWR and CRQ question sets were 2.96 and 2.72 respectively.

The five content areas were examined individually in a number of ways for differences across the two test formats. These included differences in proportion correct, proportion providing any answer at all, proportion providing an incorrect answer that was from among the multiple-choice options, and in patterns of particular incorrect answers. The first three of these analyses involved combining the data from question sets 4 and 5, whereas the last involved separate analyses of each set.

First, differences in proportion correct were examined across the content areas by combining the data from sets 4 and 5. Table 3(A) shows the difficulty levels of the two formats across the five content areas. The results reveal a significant difference, favoring the multiple-choice condition, for only one content area, crystal structure ($p<0.05$).

A second examination of the data from sets 4 and 5 concerns the proportion of students who failed to provide any answer at all. Table 3(B) contains the proportions of students who did not provide an answer under the two conditions. For all content areas, fewer students provided no answer in the CRQ format; the difference was significant only for the buffers questions, which were also by far the most difficult of the questions (Table 3(A)).

The third examination, reported in Table 3(C), concerns the proportion of wrong answers given by students in CRQ conditions that were not offered as multiple-choice options. What these data reveal is the success of the multiple-

Table 3
Differences in Proportions, Questions Sets 4 and 5 Combined (+ favoring MCQWR)

	MCQWR	*CRQ*	Δ
(A) Differences in Proportion Correct			
Gas Laws	.65	.60	+.05
Kinetics	.71	.74	−.03
Crystal Structure	.63	.51	+.12[1]
Stoichiometry	.56	.54	+.02
Buffers	.43	.34	+.09
(B) Differences in Proportion Not Answering			
Gas Laws	.02	.04	−.02
Kinetics	.01	.02	−.01
Crystal Structure	.02	.09	−.07
Stoichiometry	.02	.08	−.06
Buffers	.07	.28	−.21[1]
(C) Differences in Proportion of Incorrect Answers Not Provided as M-C Options			
Gas Laws	.00	.65	−.65[1]
Kinetics	.00	.38	−.38[1]
Crystal Structure	.00	.69	−.69[1]
Stoichiometry	.00	.70	−.70[1]
Buffers	.00	.71	−.71[1]

[1] $p<0.05$.

choice distracters in capturing the variety of likely wrong responses. Keeping in mind that these distracters were carefully constructed using archived constructed-response tests as sources of potential distracters, Table 3(C) reveals that in only one content area, kinetics, was this particularly successful. However, because the kinetics questions had the highest proportion of correct answers and hence were the easiest of the content areas, one would expect there to be less variation in the incorrect responses. In the remaining four content areas a great variety of other wrong answers were given by the students.

The fourth analysis examined the patterns across the two formats of incorrect responses that were provided as multiple-choice distracters. This analysis had to be carried out on each question set separately and did not take into account the "other" wrong responses. For questions in both sets, the Chi-square goodness of fit yielded nonsignificant values in all cases except one, indicating a reasonably good fit between the patterns of MCQWR errors and the patterns of those same errors in CRQ answers. Because both data sets yielded similar patterns, only the data for set 4 are reported here. Table 4 provides response percent data for all answer options.

The written responses to question sets 4 and 5 were also analyzed for similarities and differences in the solution path used to generate the final answer. The intent of this analysis was to examine differences in the solution paths leading to correct answers for multiple-choice and equivalent con-

Table 4
Response Choice Data (percent) for MCQWR versus CRQ (question set 4)

	Answer Option						
Question	*A*	*B*	*C*	*D*	*E*	*N/A*	*Other*
Gas Laws							
MCQWR	13	62*	14	8	1	2	0
CRQ	4	62*	4	0	0	2	28
Kinetics							
MCQWR	74*	4	9	7	6	0	0
CRQ	72*	1	6	2	4	1	14
Crystal Structure							
MCQWR	16	8	5	63*	7	1	0
CRQ	5	4	2	54*	1	7	27
Stoichiometry							
MCQWR	7	5	5	16	65*	2	0
CRQ	0	0	0	4	67*	7	22
Buffers							
MCQWR	6	15	13	40*	18	8	0
CRQ	0	4	4	34*	1	30	27

*correct answer.
N/A: no answer provided.
Other: answers other than those provided in the options.

structed-response questions and to compare across question format the solution paths that generated incorrect answers. Analysis of solution paths was possible only for those answers that were accompanied by a complete solution. Therefore, the written responses were first sorted for complete versus incomplete responses and then for correct versus incorrect final answers. Then the grouped response data were analyzed for similarities and differences in the solution paths used. Because this analysis was content-dependent, the response data had to be analyzed question by question. Results of this analysis were consistent for all question-content areas. For the sake of brevity, only one of the five question-content analyses, that for the partial pressures question, is presented in Table 5. For responses that produced a correct final answer to this question (5(B)), there was a marked similarity across question format in the type of solution path and frequency with which it was employed. For responses that generated an incorrect answer (5(C)), the type and frequency of strategies used were also similar across question formats.

Table 5
Analysis of Solution Strategies used for Partial Pressures Question

	Proportion of Answers	
A. Solution Responses for All Answers	*MCQWR (n=260)*	*CRQ (n=259)*
no answer	0.01	0.03
incomplete, incorrect/no answer	0.07	0.01
incomplete, correct	0.03	0.01
incomplete, incorrect	0.26	0.36
complete, correct	0.63	0.59
B. Strategies for Complete, Correct Answers	*MCQWR (n=164)*	*CRQ (n=153)*
added moles	0.46	0.42
added partial pressures	0.46	0.56
added volumes	0.01	0.01
made math errors, still got correct answer	0.04	0.01
illogical solution path	0.03	0.00
C. Strategies for Complete, Incorrect Answers	*MCQWR (n=68)*	*CRQ (n=95)*
used K_p formula	0.01	0.02
used Combined Gas Law	0.00	0.08
added weights not moles	0.29	0.30
made incorrect assumption	0.09	0.09
used incorrect molecular weight	0.34	0.28
solved for one partial pressure only	0.09	0.07
made math error	0.12	0.16
illogical	0.06	0.00

Analysis of Oral Reports

Coding scheme for oral protocols. The coding scheme used in this study to analyze the oral reports evolved from pilot interviews and the work of Hayes (1981) and Nurrenbern (1979). The final coding scheme consisted of four categories, which are described in the following sections A to D. Each category was defined by a series of typical behaviors. Because the categories were not mutually exclusive, the coding method did allow for a single statement to be classified in more than one category.

A. Problem representation. Information-processing theories of problem solving emphasize the importance of the generation of a problem representation or problem space (Gick, 1986; Greeno, 1978a, 1978b; Hayes, 1981; Newell & Simon, 1972). When constructing a representation of the problem the solver attempts to understand the problem by extracting the given information and connecting it to existing knowledge so as to identify the gap that must be crossed. In this study types of behavior that were classified as problem representation included rereading the problem, drawing a diagram, and defining the problem goals.

B. Strategy development. Evidence of strategy development was defined as any explicit reference through word or action to strategies to solve a problem or to steps of a solution path or formula.

C. Dominant strategy. A dominant problem solving strategy was identified for each subject on each problem. The strategy was based on the major sequence of processes used by the subject during the problem solving event. Where more than one strategy was used, both strategies were identified and labeled according to the sequence in which they were employed. The categories of strategies were derived from the model used by Nurrenbern (1979) and were modified to meet the goals of this investigation.

D. Evaluation processes. Each protocol was scored for types of evaluation behavior used to verify or assess the problem solution.

Coding reliability. The reliability of the coding scheme was assessed for both reproducibility and stability. Reproducibility or intercoder reliability was determined by comparing the protocol scoring decisions of two independent judges (one of the investigators and a third party) on four oral protocols. The intercoder agreement was 87.8% (123 out of 140 decisions made from the four protocols). Stability over time was determined by comparing coding decisions over an interval of two months; agreement was found on 89.1% of the 175 decisions made.

Results of Protocol Analysis

A data base of coded statements per subject-protocol was created. Because it was difficult to quantify the amount of a particular behavior, only the presence or absence of a behavior was scored. The frequencies for each behavior in a problem solving category were summed across protocols for each question format and question-content area. Then the frequencies for each behavior were summed in each question format across the five question-content areas. The frequencies for each of the four behaviors are reported in the two rightmost columns of Table 6.

Although the sample was too small to permit statistical analysis, it is interesting to note the differences in problem solving behavior across question

Table 6
Analysis of Verbal Reports (*N*=29)

	Frequency (%) MCQWR	CRQ
A. Problem Representation		
rereads problem or parts of problem	69	72
restates problem in own words	21	24
draws a diagram	10	14
makes a list	21	38
uses mnemonic notation	7	10
defines goals of problem	35	31
B. Strategy Development		
refers to algorithm/formula	83	86
refers to steps of a solution	59	62
recalls related concept	14	14
recalls related problem	7	24
reasons inductively/deductively	14	14
exhibits propositional logic	10	0
uses proportional logic	31	14
separates parts of problem	45	45
defines no strategy	7	0
C. Dominant Strategy		
algorithmic	62	83
algorithmic-reasoning	24	17
random trial and error	17	0
systematic trial and error	0	0
working backward	0	0
D. Evaluation Processes		
routine check of manipulations	7	28
checks that solution satisfies conditions	0	14
checks by retracing steps	3	7
derives solution by another method	14	14
is the result reasonable?	28	21
compares with generally known results	0	0
expression of confidence in answer	17	28
expression of lack of confidence in answer	24	14

format. For example, in the category of problem representation (Table 6(A)), there appeared to be a difference in the frequency of list-making behavior. Students who answered a question in constructed-response format tended to make lists more frequently than did students answering the same question in MCQWR format. This difference was more evident in questions that contained a considerable volume of data (e.g., the gas laws and stoichiometry questions). Also striking about this first section of Table 6 is that the predominant method of representing a problem, regardless of format, was to simply reread it.

In the development of strategy (Table 6(B)), students who solved MCQWR questions tended to use propositional and proportional logic more frequently than did students who answered the same questions in constructed-response format. This difference was traced to the kinetics question where proportional and propositional logic were most appropriate.

With respect to the dominant strategy used (Table 6(C)), the results indicate that most students used an algorithmic approach regardless of the question format and that none of the students employed either a systematic trial-and-error or a working-backward approach to the problems. It is interesting to note that the incidence of random trial and error, although small, was completely restricted to questions in MCQWR format.

The frequencies of different types of evaluation processes were too small to justify conclusions about any particular type (Table 6(D)). It was of some interest that the most frequently employed evaluation strategy was routine checking of manipulations. In order to get a sense of the extent to which evaluation activities were employed for the two test formats, the data for the first six categories were collapsed. A comparison of the frequencies of the collapsed categories across test format indicated a higher frequency of evaluation behavior with the constructed-response format. Also, a greater frequency of expressions of confidence was associated with constructed-response answers. Conversely, expressions of lack of confidence occurred more frequently with multiple-choice (MCQWR) questions.

Discussion

The results indicated that there was no significant difference in the success rate of students on the two question (MCQWR versus CRQ) formats as measured by mean total performance scores. The provision of multiple-choice options did not appear to enhance the ability of students to obtain the correct answer. From the written and oral reports it was apparent that students generally proceeded through a full solution process to determine the answer regardless of question format.

Analysis of the patterns of incorrect responses presented in Table 3 did reveal some format-related differences. A wider variety of errors occurred in answers to constructed-response questions. This was, however, not surprising given that for the multiple-choice questions the students were limited to four incorrect options. In fact the acid test of a multiple-choice question is its ability to provide incorrect options that comprise the majority of errors that students are likely to make in that particular problem. In this investigation the incorrect multiple-choice options were developed based on a review of the most common errors produced by similar groups of students to the same question posed in a constructed-response format. Even using this method to design the multiple-choice questions, the extent of other-answer production in the equivalent constructed-response questions varied from 38% to 71% of the group who incorrectly answered those constructed-response questions.

The question that produced the highest proportion of other-answer responses was the crystal structure question. From the oral reports, it was evident that most students had no sense of what would constitute a reasonable answer to the question. In other questions students appeared to be guided to some extent by the size of the expected answer and would reattempt the

solution if they produced what they felt was an unreasonable answer. For this question few students communicated a sense of what was reasonable. It was also apparent from the written and oral reports that many students had little understanding of the conceptual basis for solving this problem. This is consistent with the observation by Nurrenbern (1979) that students often solve chemistry problems by forcing the given information into a previously memorized formula (or in this case a formula provided on the data sheet). It should be noted that this problem was classified as an application problem and was not considered by the judges to be more challenging than the other set 4 and 5 problems.

When the answer pattern comparison was restricted to only those incorrect answers provided in the multiple-choice questions, the results indicated little difference between the two formats (see Table 4). Students appeared to make particular errors with similar frequencies for both question formats.

One question that did reveal a format-related difference in incorrect answer pattern production was the stoichiometry question. Because this question required a multilevel solution it offered more opportunity for error. Although there was no evidence of students being guided by the multiple-choice options in this question, the data suggested that different responses were elicited by the two question formats.

When the written explanations were analyzed, there appeared to be a significant difference across question format in the proportions of students who provided either no explanations or incomplete ones and yet still produced a final answer, whether it was correct or incorrect. The frequency of this type of response was significantly higher in the case of the multiple-choice question, suggesting that the incidence of risk-taking was increased for the multiple-choice format. However, it is quite possible that, in the case of the MCQWR, the students simply failed to provide the entire solution path that they used to generate the final answer. For this investigation the MCQWR and the CRQ were both scored dichotomously. However, for the purposes of student grading the mark allocation was MCQWR: 3 marks and CRQ: 10 marks. This difference in mark allocation may have affected the quantity and quality of the solution strategies provided in the written responses and possibly even the type of strategies employed. Although this problem might have been corrected or at least minimized by increasing the value of the MCQWR answers, such an adjustment would have created an atypical testing situation.

When the written answers were analyzed for differences across format in the solution paths used, the following patterns were noted.

1. There was little differentiation across test format in the types and frequencies of solution strategies employed to solve problems. Students appeared to use certain solution processes for certain types of questions with similar frequencies across test format. Most of the strategies were algorithmic in nature and followed standard application procedures as taught in class. Although evidence of strategies other than algorithmic may be difficult to discern from written work alone, the subsequent oral data indicated a high degree of consistency between the written explanations provided by the student during the exam and the oral explanations of the thought processes associated with answer production.

2. There was more differentiation across question format in the types and frequencies of errors committed during the solution process. The variety of errors committed in constructed-response answers was greater than that committed in the multiple-choice answers. Nevertheless, the frequency with which certain types of errors were made was in general similar for both formats.
3. There was limited evidence of increased use of solution truncation in the case of one multiple-choice question (the kinetics question). A number of students used the rate law equation to determine the order with respect to one reactant and then made an assumption about the value of the order with respect to the other reactant. The incidence of this response behavior was higher for the multiple-choice format than for the constructed-response format. However, the majority of students who solved the multiple-choice version of this question worked through the entire solution process in a similar fashion to those who solved the constructed-response version.

Interpretation of the oral reports revealed that the frequency of elements of problem solving behavior across test format was remarkably similar (see Table 6). Most students, regardless of format, engaged in some form of problem representation followed by an episode of strategy development. The extent and type of problem representation behavior was similar for both test formats. Only list-making appeared more frequently in the solution of constructed-response questions. Because many cognitive psychologists believe that problem representation is a critical aspect of problem solving and one that differentiates novices from experts, it is encouraging for multiple-choice proponents to note that for questions of this cognitive level and for this domain problem representation behavior was not differentiated across test format.

The extent and type of strategy development behavior in which students engaged did not differ dramatically across format. The most common type of strategy employed was an algorithmic one. It was surprising to note that the frequency of use of propositional and proportional logic was slightly greater in the case of the multiple-choice format. Logical reasoning strategies such as these are generally considered to belong to the higher-order thinking skills. Therefore, it may be somewhat surprising that these reasoning strategies appeared more frequently for multiple-choice questions, because it is often alleged that students tend to rely on lower-level thinking processes to answer multiple-choice questions. However, the sample size used in this study is too small to support a definitive conclusion.

There was no evidence of the working-backward strategy for either test format. Using this strategy the solver first identifies the projected goal and then works backward from that goal to develop a strategy. It is thought that working backward is easier with a multiple-choice question because the projected outcome is contained in a limited set of options. However, the findings of this study indicated no incidence of its use. This evidence is encouraging news for multiple-choice proponents. If we assume that working backward is an undesirable problem solving strategy, then its employment to answer test questions should be discouraged. Nevertheless, a much larger sample would be required to address this issue definitively.

With respect to the extent of evaluation processes employed during the test answer generation, the data indicated that evaluation processes were used more frequently to answer constructed-response questions. This may have been attributable to the greater mark allocation for CRQ. On the other hand, it may have been that the provision of answer options in the MCQ (and MCQWR) leads the solver to believe that an evaluation step is unnecessary. If this had been the case, one might have expected a higher incidence of expression of confidence associated with the multiple answers as opposed to the constructed-response answers. Yet the results showed the reverse in terms of frequency of expressions of confidence or lack thereof. Once again, the sample size is not large enough to provide a definite answer.

It is acknowledged that the time lapse between the written tests and the oral reports may have affected the quality and accuracy of the oral data. Ericsson and Simon (1980) indicate that oral reporting about a task becomes more variable as the time between the task and the reporting event increases. In this study the students were encouraged to simply report all thoughts that had transpired during the test task, and from the protocols it is clear that some students were more successful at doing so than others. In fact, what was vocalized undoubtedly involved some reporting and some reconstructing. The extent of new constructions was minimized by providing students with copies of their written test responses. Furthermore, a comparison of the reports generated during the week following the examination with those produced one month after the exam revealed no difference in level of detail or consistency with the written test answers.

Conclusions

This investigation has revealed that for questions that require some idea and data manipulation the solution processes employed to answer these questions in the two formats are remarkably similar for first-year university students in science and science-related programs of study. The findings support the recent hypothesis by Traub (1992) that questions that pose scientific or mathematical problems appear to be impervious to format effects. Yet these findings also indicate that at a finer level of examination there appear to be differences in elements of the problem solving behavior associated with the two test formats.

Given the pervasiveness of the written test as an evaluation instrument and the extent to which educators depend on the instrument for myriad decisions, the findings of this study have significant implications for educators and students as well as for researchers.

Educators may find the results of this study helpful in selecting student evaluation instruments. The findings suggest that for first-year university students pursuing science-related programs of study, response strategies for the two most common test formats are remarkably similar for the types of questions used in this study. It should be noted that these questions all involved multistep calculations and in that sense forced the students to proceed through the same set of steps to achieve the correct answer. For these kinds of questions the findings indicate that both test formats may be equally appropriate for assessing problem solving ability for this type and level of student. Educators should be cautioned, however, not to extend this interpretation to questions that test strictly fact recall or other types of reasoning strategies. Indeed, re-

search to date suggests that, for these kinds of tasks, format does make a difference (Traub, 1992). Future studies that incorporate questions dealing with fact recall and other types of tasks as well as those that examine other student samples could enhance the understanding of test response behavior.

Educators may also benefit from the insight provided by this study into the types of solution strategies employed to answer certain kinds of questions and the patterns and frequencies of errors committed. The high incidence of use of algorithmic strategies coupled with the low incidence of use of proportional and propositional reasoning strategies found in this sample of first-year university chemistry students should give chemistry educators cause for concern. It is a sobering reminder of the chasm that exists between problem solving success in a test situation and deeper conceptual understanding.

The results also have potential implications for students of chemistry. It has been suggested that students study differently depending on the type of test they expect (Loftus, 1971; Traub & MacRury, 1990; Treversky, 1973). However, according to the results of this investigation, for the kinds of test questions considered here students might be better advised to study in the same manner regardless of the test format expected, because it appears that they tend to use the same strategies to answer questions in both formats.

Finally, this investigation contributes to a better understanding of the effect of test format in the evaluation of specific traits. It has only recently been recognized that test format effects are domain- and task-specific, and investigators have now begun to define these effects for particular disciplines and types of tasks. The present finding of minimal format differences for chemistry undergraduates and for this task type provides an additional piece in the increasingly complex puzzle of trait equivalence of test formats.

References

Ackerman, T.A., & Smith, P.L. (1988). A comparison of the information provided by essay, multiple choice, and free response writing tests. *Applied Psychological Measurement, 12,* 117-128.

Bennett, R.E., Rock, D.A., & Wang, M. (1991). Equivalence of free response and multiple choice items. *Journal of Educational Measurement, 28,* 77-92.

Bennett, R.E., Rock, D.A., Braun, H.I., Frye, D., Spohrer, J.C., & Soloway, E. (1990). The relationship of expert-scored constrained free response items to multiple-choice and open-ended items. *Applied Psychological Measurement, 14,* 151-162.

Birenbaum, M., & Tatsuoka, K.K. (1987). Open-ended versus multiple-choice response formats—It does make a difference for diagnostic purposes. *Applied Psychological Measurement, 11,* 387-395.

Bloom, B.S. (Ed.). (1956). *Taxonomy of educational objectives, Handbook 1: Cognitive domain.* New York: Davis McKay.

Bloom, B.S., & Broder, L.J. (1950). Problem solving processes of college students. *Supplementary Educational Monograph, 73.* Chicago, IL: University of Chicago Press.

Connolly, J.A., & Wantman, M.J. (1964). An exploration of oral reasoning skills in responding to objective test items. *Journal of Educational Measurement, 1,* 59-64.

Ericsson, K.A., & Simon, H.A. (1980). Verbal reports as data. *Psychological Review, 87,* 215-251.

Farr, R., Pritchard, R., & Smitten, B. (1990). A description of what happens when an examinee takes a multiple-choice reading comprehension test. *Journal of Educational Measurement, 27,* 209-226.

Gick, M.L. (1986). Problem solving strategies. *Educational Psychologist, 21,* 99-120.

Glass, G.V., & Hopkins, K.D. (1984). *Statistical methods in educational psychology* (2nd ed.) Englewood Cliffs, NJ: Prentice-Hall.

Greeno, J. (1978a). A study of problem solving. In R. Glaser (Ed.), *Advances in instructional psychology* (vol. 1, pp. 13-75). Hillsdale, NJ: Erlbaum.

Greeno, J. (1978b). Understanding and procedural knowledge in mathematics instruction. *Educational Psychologist, 12*, 262-283.

Hayes, J.R. (1981). *The complete problem solver*. Philadephia, PA: Franklin Institute Press.

Hayes, J.R., Flower, L., Schriver, K., Stratman, J., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics* (vol. 2, pp. 176-240). Cambridge, UK: Press Syndicate of the University of Cambridge.

Kropp, R.P. (1956). The relationship between process and correct item-response. *Journal of Educational Research, 49*, 385-388.

Loftus, G.R. (1971). Comparison of recognition and recall in a continuous memory task. *Journal of Experimental Psychology, 91*, 220-226.

Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Norris, S.P. (1992). A demonstration of the use of verbal reports of thinking in multiple choice critical thinking test design. *Alberta Journal of Educational Research, 38*, 155-176.

Nurrenbern, S.C. (1979). Problem solving behaviors of concrete and formal operational high school chemistry students when solving chemistry problems requiring formal reasoning skills. *Dissertation Abstracts International, 40(4986A), 167*. (University Microfilm No. DEM 80-05926)

Snow, R.E. (1980). Aptitude, learning, and instruction. *Cognitive Process Analyses of Aptitude, 1*, 27-64.

Traub, R.E. (1992). On the equivalence of traits assessed by multiple choice and constructed response tests. In R. Bennett & W. Ward (Eds.), *Construction versus choice in cognitive measurement* (pp. 29-44). Hillsdale, NJ: Erlbaum.

Traub, R.E., & MacRury, K. (1990). Multiple choice versus free response in the testing of scholastic achievement (Original in German). In K. Ingenkamp & R.S. Jäger (Eds.), *Tests und Trends 8: Jahrbuch der Pädagogischen Diagnostik* (pp. 128-159). Weinheim & Basel: Beltz Verlag.

Treversky, B. (1973). Encoding processes in recognition and recall. *Cognitive Psychology, 5*, 275-287.

van den Bergh, H. (1990). On the construct validity of multiple choice items for reading comprehension. *Applied Psychological Measurement, 14*, 1-12.

Ward, W.C. (1982). A comparison of free response and multiple choice forms of verbal aptitude tests. *Applied Psychological Measurement, 6*, 1-11.

Ward, W.C., Frederickson, N., & Carlson, S.B. (1980). Construct validity of free-response and machine-scorable forms of a test. *Journal of Educational Measurement, 17*, 11-29.

The Alberta Journal of Educational Research *Vol. XLI, No. 1, March 1995, 36-62*

Jeremy Hull
Four Directions Consulting Group

Ron Phillips
Phillips and Associates

and

Eleoussa Polyzoi
University of Winnipeg

Indian Control and the Delivery of Special Education Services to Students in Band-operated Schools in Manitoba

The article addresses the need for special education programs for Indian students in Manitoba and relates this to a review of special education program delivery and supports in band-operated schools in the province. Although indicators of need suggest that Indian students may have greater and more complex special education needs than the general provincial population, until recently few Indian students living on reserves in Manitoba have been identified as having special needs. The low identification rate may be related to both lack of resources and mistrust within Indian schools and communities of special education assessment methods and programs. A review of special education programs in five band-operated schools shows that there are problems in the development, monitoring, and operation of special education programs in these schools. Although in theory the financial resources available for Indian special education programs compare favorably with those available in provincial schools, in practice Indian schools are functioning in isolation without the benefit of a regional or provincial system of specialist, planning, and monitoring supports. Five interrelated steps are proposed that could improve special education services in Manitoba's Indian schools.

Cet article traite du besoin d'avoir des programmes d'orthopédagogie pour les élèves autochtones du Manitoba et rattache se besoin à la révision de la présentation du programme d'orthopédagogie et des programmes d'appui dans les écoles gérées par des groupes amérindiens dans la province. Même si les indicateurs des besoins suggèrent que les élèves autochtones auraient des besoins orthopédagogiques plus nombreux et plus complexes que ceux de la population provinciale générale, ce n'est que récemment que les élèves autochtones habitant les réserves au Manitoba ont été identifiés comme ayant des besoins spéciaux. Le taux diminué de cette identification est possiblement relié à la pénurie de ressources orthopédagogiques et à une certaine méfiance qui règne dans les communautés amérindiennes et les écoles autochtones vis-à-vis les méthodes d'évaluation et des programmes orthopédagogiques. L'étude de programmes en orthopédagogie dans cinq écoles gérées par les autochtones

---

Jeremy Hull has worked as an education and research consultant in Saskatchewan and Manitoba since 1973 and is currently a partner in a Winnipeg consulting company.
Ron Phillips is a private consultant specializing in Native education. He recently completed his doctorate in special education at the University of Arizona.
Eleossa Polyzoi is an associate professor of education. She has been involved in several special education program evaluations and is currently evaluating special education policy and practice in the Czech Republic.

démontre qu'il y a des problèmes qui existent en ce qui concerne le développement, la progression, et l'opération de ces programmes orthopédagogiques dans ces écoles. Quoiqu'en théorie les ressources financières apparentes disponibles pour les programmes orthopédagogiques dans les écoles autochtones comparent bien avec les ressources disponibles dans les écoles provinciales, les écoles autochtones fonctionnent en réalité dans un isolement sans recours aux bénéfices d'un système de spécialistes régionaux ou provinciaux, à la planification, et sans appui. Cinq étapes entrelacées sont proposées qui pourraient améliorer les services en orthopédagogie dans les écoles autochtones du Manitoba.

Over the past several years increasing concern has been expressed over the adequacy of special education service delivery for students attending reserve schools. Because of the isolation of many reserve schools and the lack of local specialists and support services, Indian[1] students are less likely than other students to be provided with appropriate programs to meet their special needs. Studies to date have shown that Indian students have a higher incidence of age-grade deceleration and a higher rate of early withdrawal from school compared with other students. Special educational services are rarely available in the students' native language. Results of standardized achievement tests also suggest that Indian students' skills in mathematics and language arts are much lower on average than those of Canadian students generally, and the gap between Indian achievement levels and those of the general population becomes greater as the grade level increases. Socioeconomic conditions in reserve communities also have a negative impact on students. Because of their families' social and economic circumstances, Indian students are at greater risk of leaving school before reaching grade 12. Again, this problem is compounded for Indian children with special needs.

The purpose of this article is to review the need for special education programs for Indian students in Manitoba and the current state of special education program delivery in band-operated schools. First, an overview is provided of the way in which the concept of *Indian control of Indian education* has been defined and how special education has been viewed within that definition. Second, various aspects of the need for special education services are discussed, including enrollments, success indicators, and socioeconomic conditions. Third, the delivery of special education programs in five band-operated schools is described. This is followed by a description of the current organizational structures supporting the delivery of special education to band-operated schools. This discussion focuses on the system for providing funding and support services to local schools, using the provincial system in Manitoba for comparison. The article concludes with a number of recommendations for improving the delivery of special education in band schools.

Indian Control of Indian Education: An Overview

The concept of Indian control of Indian education derives from the spirit of Aboriginal self-determination and from the various treaties that were signed between Aboriginal groups and the Crown. In exchange for rights to Aboriginal lands, the treaties promised among other things to provide schools on reserves when requested to do so (Morris, 1862). However, since the 1970s Indian organizations and leaders have been critical of the manner in which education has been provided to Indian communities. Criticism has focused on

the poor success of Indian students and lack of participation of Indian people in the decision making structures affecting Indian schooling.

In 1972 the National Indian Brotherhood (NIB, 1972) developed a policy paper on *Indian Control of Indian Education*. This paper, which was subsequently accepted in principle by the Minister of Indian Affairs and Northern Development, became the basis for developments in Indian education over the next 20 years. It dealt with philosophy of education, jurisdiction, programs, teachers, and facilities. However, the issue of special education was not specifically addressed.

The NIB (1972) paper strongly advocated that band councils have control of decisions and that parents have a prominent role in decision making, but the paper did not argue that band-administered education is always preferable to education in provincial schools. In particular, there was concern that the federal government had been making agreements with provincial school authorities to provide for Indian education without the consent or participation of the bands concerned. It was proposed that each band decide whether to operate its education program directly or to enter into an agreement with a provincial school board.

The NIB (1972) paper emphasized that the schools should reflect Indian values and should incorporate Indian history and language in the curriculum. In addition, the paper argued for the provision of nursery and kindergarten programs in the schools, expansion of secondary, postsecondary, and adult education programs in Indian communities, and provision of drug and alcohol education and prevention programs. The paper also called for Indian controlled cultural education centers, the training of Indian teachers and counselors, employment of Indian paraprofessionals, and improvement of educational facilities.

Since the NIB (1972) paper was first produced, there has been a dramatic shift to locally administered reserve schools. In addition, nursery and kindergarten programs offered for four- and five-year-olds as part of the school program are now standard, many Indian teachers have been trained and are now teaching on reserves, and much greater participation rates of Indian students in postsecondary programs have been reported. For example, the number of Indian postsecondary students in Canada doubled between 1985 and 1991, reaching more than 21,000 in 1991-1992 (INAC, 1992a, p. 39). Many of these postsecondary students have become teachers and are now teaching in reserve schools. As of 1986, three quarters of those employed in education on reserves in Canada were registered Indians (McBride, Gagne, & Atwell, 1990, p. 5). Most Indian schools now offer some type of Native studies program and many provide for Aboriginal languages as a medium of instruction or as a subject area. Many bands have established elected or appointed education authorities that have responsibility for administering the band's educational programs. And it is now the policy of INAC that agreements with provincial schools for the provision of education to Indian students must always involve the affected bands (McKnight, 1988).

Despite these advances, some feel that Indian educational authorities are simply administering a non-Indian school system. For example, Koens (1989) argues that immediately after an Indian school comes under local control, there

is a burst of enthusiasm leading to a superficial increase in symbolic Indian cultural content such as crafts and dancing. However, when these changes fail to reduce chronic student absenteeism and other educational problems, the effort to increase Indian content in the school is typically allowed to evaporate, and the school program becomes a "pale reflection of the provincial curriculum" (p. 37).

In 1988 the Assembly of First Nations (AFN) published a major policy review of Indian education. This was an updated and much more detailed statement than the NIB (1972) paper, but the central themes were much the same. It emphasized preservation of languages and culture, instruction in Indian values, preparation for "total living," that is, development of skills and abilities to function in a variety of occupations and cultural contexts, and the preeminence of the authority of the local band council in making educational policy.

The review identified a number of problems with the use of standardized tests in First Nations schools, noting that achievement tests are widely felt to be culturally biased and that they fail to take account of differences among provinces in curriculum. Biased tests were viewed as being responsible for channeling Indian students into inappropriate remedial, occupational, and special education programs. Moreover, the review found that many band schools did not have the capability of undertaking diagnostic testing to screen for students with learning disabilities or to identify gifted students (AFN, 1988, p. 80).

On the other hand, the AFN review found that most First Nations communities surveyed were concerned about the lack of appropriate special education services and funding. The review found that "special programs with intensive remediation" are required for the teaching of basic skills, for speech and language development, for development of learning skills, and for the education of multihandicapped and developmentally delayed children (AFN, 1988, p. 88). Concerns expressed about special education included:

- classification of Indian students as having special needs primarily as a means of obtaining additional funding;
- lack of student assessment practices;
- failure to obtain permission of parents before placing students in modified programs;
- lack of local control over placement decisions;
- failure to address the needs of gifted students;
- failure to screen for health problems that may result in learning disabilities; and
- lack of accountability of provincial education systems to parents of First Nations students in special education programs.

Many of the issues that affected Indian education 20 years ago are not yet resolved. Although band councils and Indian education authorities now have much more authority in administering educational programs than they used to, these authorities have not fully assumed control over program planning or have not had the resources to implement the programs they desire.

The area of special education is a case in point. As is argued below, there is reason to think that special education needs in Aboriginal communities might

be greater than those in other communities. However, the AFN review makes it clear that there is a great deal of distrust of the area of special education in Aboriginal communities and schools. The fear is partly that Aboriginal children are being inaccurately labeled as having learning disabilities and are then shunted into classes that give them few opportunities to develop. Special education is also an educational specialization in which outside experts such as psychologists make decisions about the children. Local school staff and parents may feel intimidated and left out of the process.

The fundamental issue affecting the control and direction of Indian school systems, including the provision of special education services, is how control can be exercised locally when the funding comes from the federal government. Related to this is how the school administration and board are to be held accountable for their success in meeting educational needs and goals of the community. With funding provided by the federal government and the board elected by the community, the logic of accountability is broken. As Paquette (1986) notes, "The reality experienced by most aboriginal people, who are not ratepayers in a public system, is one of an educational system all of whose capital and operating revenues, and most of whose human resources as well, come from the 'outside'" (p. 21).

This general issue of control and accountability is more sharply drawn in the case of special education services. The somewhat specialized and technical nature of special education makes it a more difficult area for parents and school boards to come to grips with. For the same reason it may seem more difficult to incorporate Indian values into special education services. Despite the increase in the number of Indian teachers, few of these teachers have specialist training related to special education. The continuing distrust of the role and value of special education programs and assessment methods on the one hand, and the problem of meeting the educational needs of the students on the other, present Indian parents and educators with a dilemma: Should Indian schools adopt the non-Indian system that they may distrust? And if not, what new system of dealing with student's special educational needs will be used?

Underlying this dilemma is the question, what are the special educational needs of Indian students? The following section attempts to provide an overview of this question.

The Special Educational Needs of Indian Students

The information on Indian enrollments, success indicators, and socioeconomic characteristics given in this section is intended to provide the necessary background for understanding the need for special education programs for Indian students. The statistics given below describe Indian students under federal jurisdiction whose permanent residence is on reserve, even though they may be attending a provincial school and may be living away from home temporarily. These statistics do not include Indian students whose normal place of residence is off reserve.

Enrollment Trends

Between 1975 and 1991 the number of registered Indian students living on reserves in Canada increased by approximately 35%, from about 72,000 to 97,000 (INAC, 1992a, p. 43). Indian students are enrolled in three different

types of schools: Schools operated directly by the federal Department of Indian Affairs (federal schools), schools operated by local school boards falling under provincial jurisdiction (provincial schools), and schools operated by band governments or education authorities, including tribal councils (band schools).

In 1975 41% of Indian students in Canada attended federal schools, but by 1991 this proportion had fallen to 6%. Over the same period the proportion of Indian students attending provincial schools fell from 53% to 45%, and the proportion attending band schools increased from 4% to 47% (INAC, 1992a, p. 43). In effect, most of the Indian student population in federal schools has been transferred to band jurisdiction, as local bands assumed control of formerly federal schools. The proportion of Indian students attending provincial schools has fallen slightly, reflecting the tendency of band-controlled schools to increase the number of grades offered on reserve rather than sending students out to provincial schools.

An even greater shift of Indian enrollments into band-operated schools has taken place in Manitoba over the past 15 years (see Figure 1). Most of the federal schools in Manitoba were transferred to band control during the 1980s and 1990s. As of September 1993, of the 45 reserve school systems in Manitoba only seven were federally operated and all of these were slated to come under local control by 1994. As can be seen in Figure 1, as of 1992 band school enrollments were up to about 12,000 and federal school enrollments were down to about 1,500. Indian enrollments in band schools now make up more than two thirds of all Indian students residing on reserves in Manitoba.

The number of Manitoba Indian students attending provincial schools, has remained constant at about 4,000 students for the past 12 years. These students may be living on reserve and attending a nearby provincial school or they may be temporarily living off reserve in private homes or in residences in order to attend school. A small number of reserve students attend institutions designed to meet special needs, such as the Manitoba School for the Deaf in Winnipeg or institutions designed for those who have severe behavior problems, such as Knowles Centre, also in Winnipeg. About half of the Indian students attending

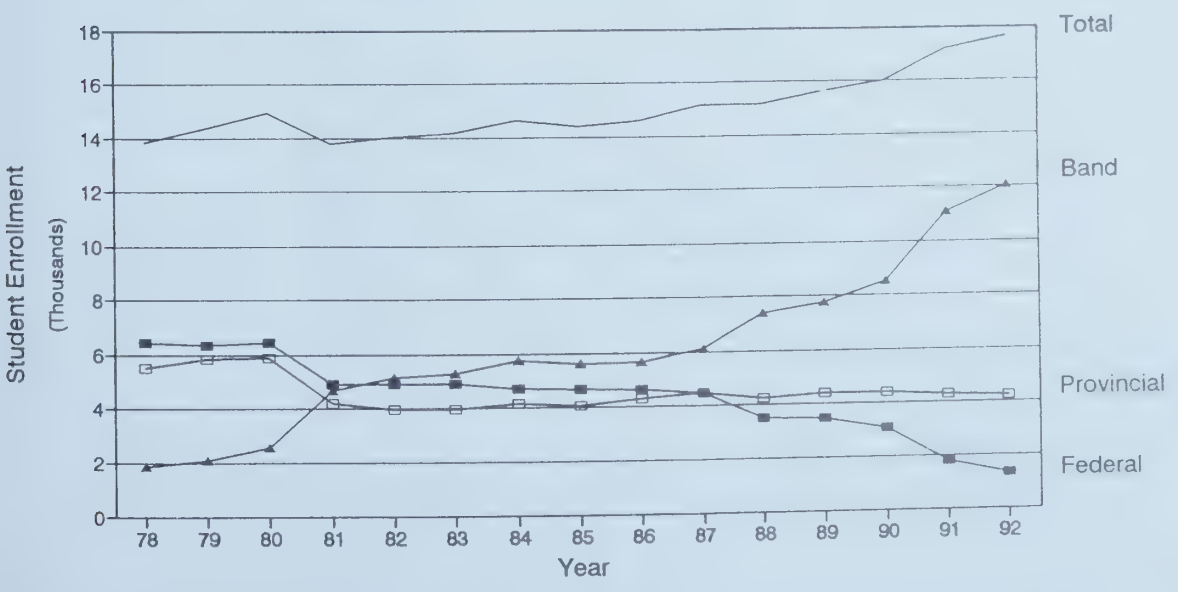


Figure 1. Enrollment of Indian students living on reserves by school type, Manitoba, 1978-1992 (source: unpublished Nominal Roll data).

provincial schools are enrolled in the Frontier School Division in the northern part of the province. More than one third of the Indian students attending provincial schools are enrolled at the high school level (unpublished 1992 Nominal Roll data provided by INAC, Manitoba Region).

Age-grade Deceleration

One indicator of the extent to which Indian students have special needs is the proportion of Indian students behind the expected grade level for their age group, referred to as age-grade deceleration.

Figure 2 compares the number of students of a given age who are one or more years behind the expected grade level for their age group in the three different school types during the 1986-1989 period. By age 12 more than half of the Indian students attending band and federal schools are decelerated by at least one year. The percentage increases dramatically in older students. This situation contrasts sharply with the degree of age-grade deceleration among the general student population of Manitoba, where it is not considered to be an issue.

The figure shows that there are slightly fewer students behind the expected grade in the provincial schools than in the band and federal schools. There are a number of possible explanations for this, such as differences in the social and educational expectations in different types of schools, differences between Indian students who choose to attend provincial schools as opposed to federal or band schools, and differences in schools' student promotion policies. To the best of our knowledge, no research has been done to date that provides a clear explanation for this difference.

At age 15 there is a noticeable increase in the proportion of students in federal schools who are behind the expected grade, and a decrease in the proportion of students who are behind in provincial schools. This may reflect selective transfers of more successful students from band or federal schools to

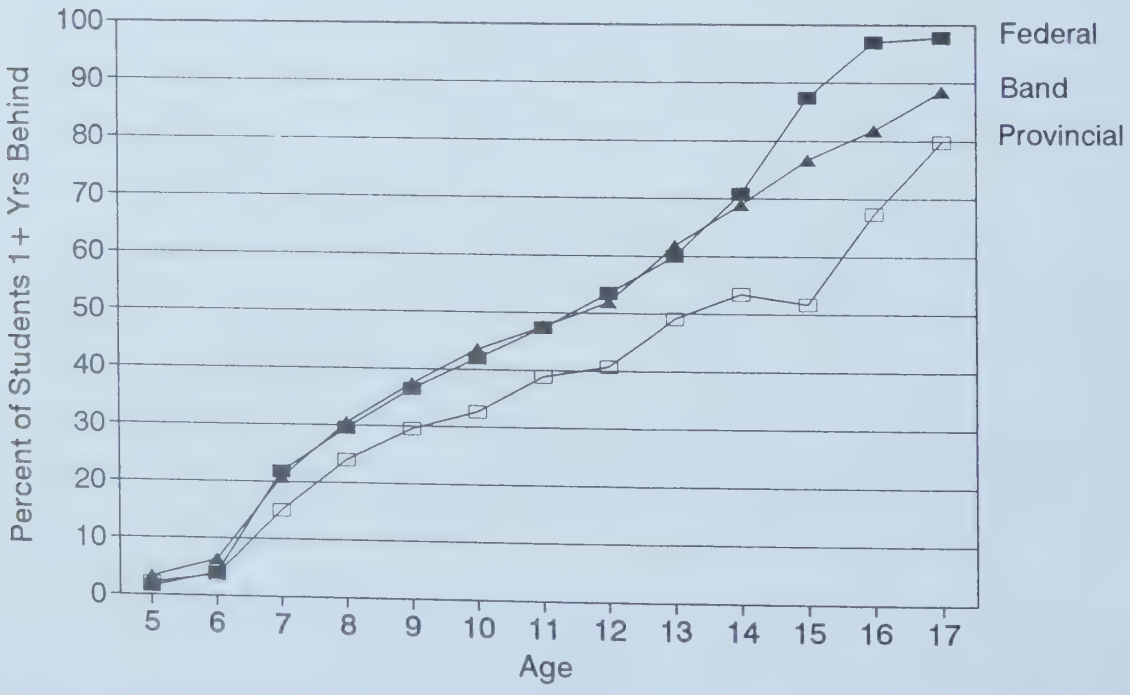


Figure 2. Manitoba Indian students behind grade by age and school type, 1986-1989 averages (source: unpublished Nominal Roll data).

provincial schools at about grade 9 or 10. It is therefore possible that those who remain in the reserve schools would be more likely to be behind the expected grade level and also more likely to withdraw prior to completing high school (personal communication, Director of Education of a rural band-operated school system, January 26, 1994).

Enrollment/Withdrawal Rates

A second possible indicator of special needs is the enrollment rate. This is defined as the percentage of students of a particular age group who are enrolled. It is, therefore, the converse of the dropout rate. An enrollment rate of 75%, for example, would suggest that 75% of the students of a given age are enrolled in school and that 25% have left school. The enrollment rates of Indian students are lower than those of the general provincial student population, but the gap between the two groups is much smaller than it has been in the past. Figure 3 compares enrollment rates among all on-reserve Indian students in Manitoba (regardless of the type of school they attend) with the enrollment rates of the general provincial student population.

As this figure illustrates, Indian enrollment rates decline rapidly after the age of 14 and are between 15% and 20% lower than provincial rates among 15-, 16- and 17-year-old students. For older students, Indian enrollment rates are higher than provincial averages, but this is partly because students in the provincial system tend to graduate from high school at an earlier age than do Indian students.

A number of studies of Indian student withdrawal rates were conducted in the 1970s. These studies reported average dropout rates across all grades of between 15% and 45% per year depending on the study (Stevens, 1982). Stevens reported that about 34% of Indian students in federal and provincial

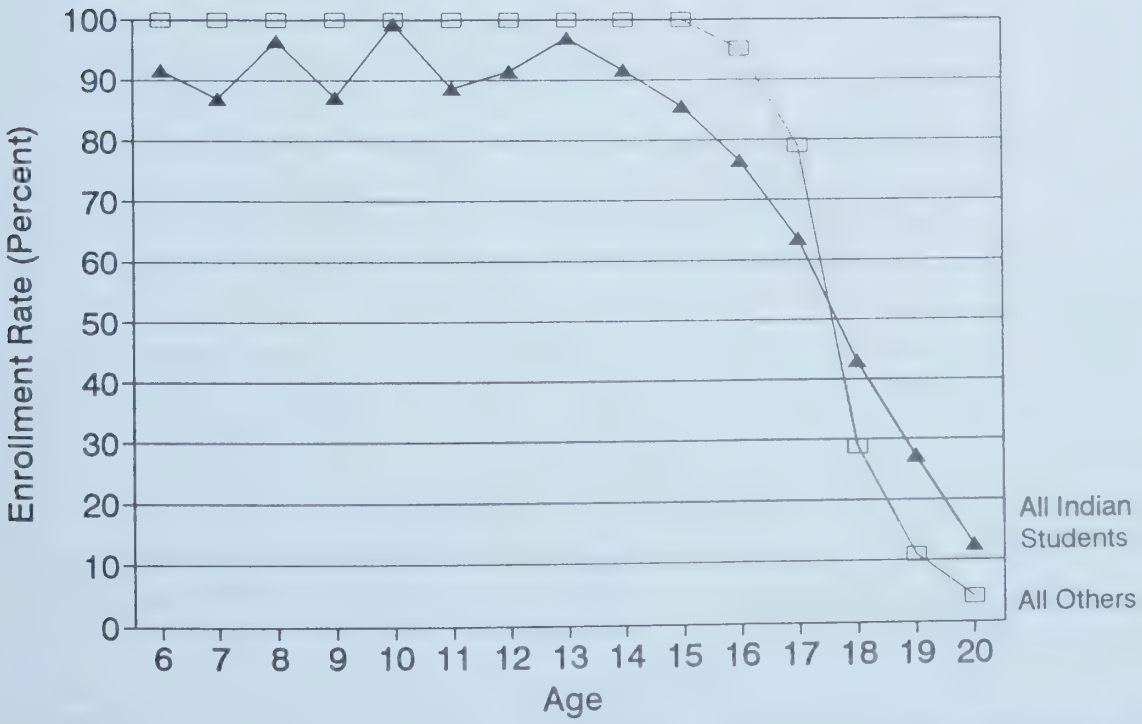


Figure 3. Manitoba K-12 enrollment rates, 1989-1990, Indian students and others, by age (source: Education in Canada, 1989-1990 [Statistics Canada #81-229]; Nominal Roll; Indian Register).

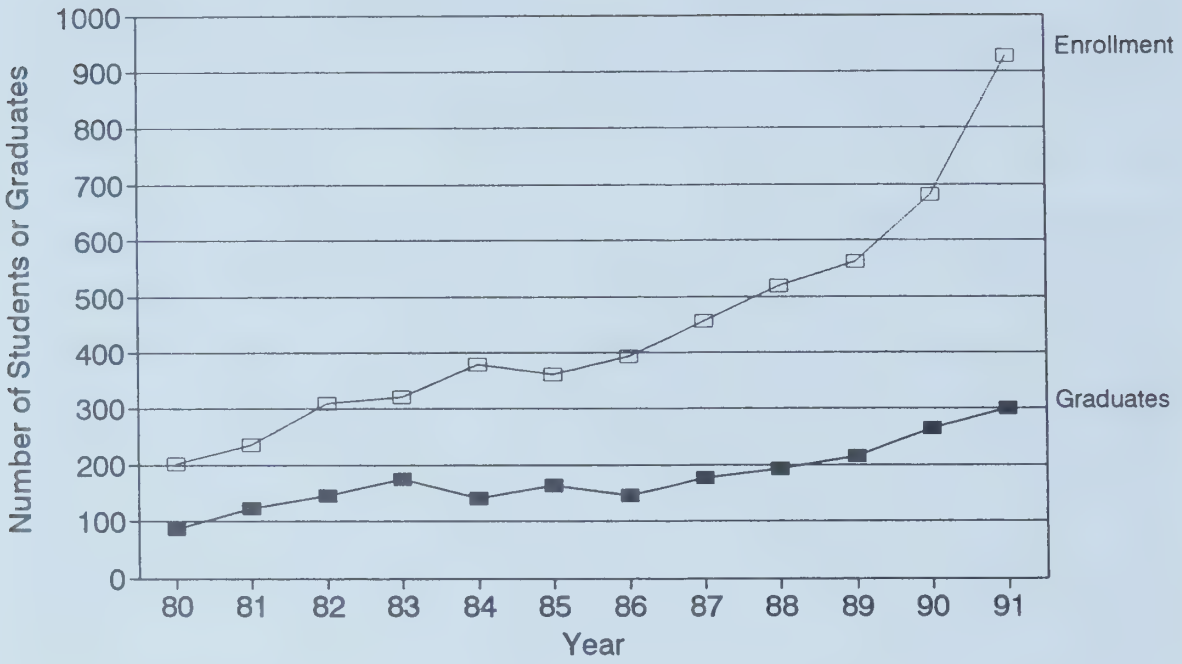


Figure 4. Indian grade 12 enrollments and graduates by school type and year, Manitoba 1980-1991 (source: unpublished data, Indian and Northern Affairs, Manitoba Regional Office).

schools in Manitoba dropped out each year during the 1977-1979 period. By comparison, the dropout rate of students in Winnipeg's inner city in 1978-1979 was 10% (p. 29).[2]

Graduation Rates

Another indicator of student success or levels of need is the graduation rate. As Figure 4 shows, both the number of Indian students enrolled in grade 12 and the number who graduate have increased over the past 12 years, but enrollments have increased more quickly than graduates. This means that the graduation rate (grade 12 graduates as a proportion of those enrolled in grade 12) has declined, from 43% in 1980 to 32% in 1991. The falling graduation rates may be a function of increased grade 12 enrollments. It may be that in earlier years only the most motivated and successful Indian students were enrolled in grade 12 and that the higher proportion of students enrolled in more recent years includes a larger proportion of students with somewhat lower motivation or academic skill levels. If true, this would be expected to lead to lower graduation rates.

Previous research found higher graduation rates among band-operated schools than among federal or provincial schools in Canada in 1980-1981 (Hull, 1990). However, in Manitoba for the period from 1980 through 1991 graduation rates in band-operated schools were lower than those in provincial or federal schools (INAC, Manitoba Region, Unpublished Nominal Roll Data, 1992b). It is possible that the lower graduation rates reflect greater success among band-operated schools in encouraging students to stay in school. This might have resulted from the combination of expansion of the number of grades offered locally (often associated with band-operated schools), selective transfers of better students to provincial schools, and the retention of students with weaker academic skills or less motivation in the band-operated schools.

Table 1
Percentages of Indian Students in Federal Schools and of All Students in Provincial Schools Who Met the Provincial Reading and Math Level Criteria Manitoba, 1980 and 1981

Grade	*Indian Students Attending Federal Schools*		*All Students Attending Provincial Schools*	
	Reading	*Math*	*Reading*	*Math*
3	44%	39%	80%	65%
6	35%	35%	62%	62%
9	42%	32%	62%	54%
12	48%	N/A[a]	61%	N/A[a]

[a]No test results reported for these groups.
(Source: Assembly of Manitoba Chiefs, 1984, pp. 85-86).

Achievement Test Results

Student success on achievement tests is another indicator of the level of need for special education services. Reading and mathematics skills of Indian students in federal schools in Manitoba were tested in 1980 and 1981 through the Manitoba Assessment Program. This program tests students in relation to provincial curriculum objectives, and was given to students in grades 3, 6, 9, and 12. As Table 1 shows, between 35% and 48% of the students in federal schools successfully met the reading level criteria, compared with 61% to 80% of Manitoba students in general. For both math scores and reading scores there was a large gap between the Indian students in federal schools, and the general provincial average (Assembly of Manitoba Chiefs, 1984, pp. 85-86).

No systematic study of Manitoba Indian students' performance on achievement tests since 1981 is available. However, a number of reserve school evaluations have been conducted over the past several years and many of these have reported achievement test results. Although various types of tests have been used in these evaluations, a number have been based on the Canadian Test of Basic Skills (CTBS). The CTBS produces a composite score that is a grade equivalent based on Canadian norms. Therefore, it has a built-in comparison with the achievement levels expected in Canadian schools. Although the CTBS has been criticized as inappropriate or culturally biased, it does at least provide a consistent reference point giving some indication of the success of students in gaining academic skills.

For the purposes of this article, evaluations of five schools in which CTBS results were reported at various grade levels were selected. The schools are not representative of all schools attended by Manitoba Indian students. They include only band or federal schools. Four are located in the northern administrative area of Manitoba. One is classified as urban, two are rural, one is remote, and one is isolated in terms of the INAC geographic classification system (INAC, 1987).

In most of these schools both the mathematics and language arts subtests of CTBS were administered. Table 2 presents these results in terms of the differences between the average grade equivalency scores of the CTBS tests and the

Table 2
CTBS Test Results Showing Average Years Behind Canadian Norms for Five Reserve Schools in Manitoba by Grade Level
Various Years from 1990 through 1993

Grade	Schools A	B	C[a]	D[b]	E	Average
1	.7	1.3	—	—	.5	.8
2	.8	—	—	—	.8	.8
3	1.1	—	.6	1.2	1.3	1.1
4	1.4	2.2	—	—	1.1	1.6
5	1.4	—	—	—	1.5	1.5
6	1.3	—	2.5	1.8	2.4	2.0
7	2.0	—	—	—	1.1	1.6
8	1.7	3.3	—	—	1.9	2.3
9	—	3.8	2.7	2.5	1.6	2.7

Note. All scores are composites of verbal and math scores except where otherwise noted.
[a]Reading comprehension subtest only.
[b]Math sub-tests only.
(Sources: Various unpublished school evaluation reports).

current grade level of the students. For example, if grade 3 students were tested at the end of November their current grade level would be 3.3 (three months into grade 3). If their average composite score was 2.5 the difference would be –.8 or eight 10ths of a year behind the norm in the skills tested.

The results shown in Table 2 indicate a trend toward greater differences in skill levels between Indian students in reserve schools and Canadian norms as the grade level increases. In other words, the longer the students are in school the farther behind they are falling. Although CTBS tests may not be sensitive to the cultural differences between Indian students living on reserves and other Canadians, they provide an indication of the success of students in obtaining skills that are valued in the job market and in Canadian society in general. Students moving from reserve schools into provincial school systems, postsecondary institutions, or the labor force are likely to face difficulties as a result of inadequate skill levels.

Socioeconomic Factors Affecting Schools

Socioeconomic conditions also provide an indication of the level of need for special educational services among the student population. The general pattern of greater educational attainment among those from higher socioeconomic status (SES) has been found to hold true for Indian students as well as others. An analysis of 1981 Census data found that as the SES index increased, the probability of 20-24-year-olds having completed high school also increased for both Indians and non-Indians (all others), and the gap between Indians and non-Indians became smaller (Hull, 1987a, p. 54). It was also found that in those Indian families with income below the poverty line, in those families where an Aboriginal language is spoken in the home, and among those living on-reserve, fewer children completed high school (Hull, 1987a, pp. 59-61). Low income

levels and high unemployment rates have been well documented among Manitoba's on-reserve Indian population (e.g., Hull, 1987b).

Special Education Enrollments in Reserve Schools

The above review of indicators suggests two things: First, there may be a higher incidence of special needs among students living on reserves than is found in other parts of Manitoba. Second, reserve students with special needs are attending school in a more difficult educational environment than other Manitoba students. Thus the difficulties they encounter because of their special educational needs are compounded by low levels of student achievement, high rates of withdrawal and age-grade deceleration, and difficult socioeconomic conditions.

Based on this reasoning we might expect a high proportion of students on reserves to be identified as having special needs. However, this has not been the case. Prior to 1986-1987, few Indian students were identified as having special needs, and the data available on special education students were limited and unreliable. Figures from the Nominal Roll during the 1978-1982 period showed that about 1.5 to 2.0% of Indian students were identified as special education students (Hull, 1987a). These figures seem unreasonably low since an estimated 10% of non-Indian students in Manitoba are believed to have special needs (Manitoba Education and Training, 1989).

In order to gain an independent estimate of the level of special needs in reserve schools, the Manitoba Indian Education Association (MIEA) conducted a survey in 1984 of 14 band-operated schools in seven Manitoba tribal councils to identify the need for special education services (Phillips & Cranwell, 1988). The survey involved personal interviews, primarily with teachers and principals, regarding the special needs of students enrolled in their schools. Results indicated that 969, or 31% of the 3,125 students surveyed, were suspected of having learning problems requiring specialized services. The survey was not intended to take the place of a thorough needs assessment, but it did provide an estimate of the total number of students who might potentially be in need of special education services.

In 1987 a report by the Island Lake Tribal Council on special education programs indicated that fewer than 1% of the students in their communities were identified as special education students on the Nominal Roll. Of this percentage, the majority tended to be classified by their medical condition. In addition, schools were often reluctant to refer students because parents were concerned over the negative effects of labelling, there were limited resources in the schools for properly assessing the child, and there was limited programming available to meet the students' needs once these needs were identified.

As recently as 1988-1989 it was found that 14 of the 32 band-operated schools in Manitoba and eight of the 13 federally operated schools were still not reporting any special needs students (Assembly of Manitoba Chiefs, 1991, pp. 82-83). These schools were located throughout Manitoba and ranged in size from seven to 747 students.

In 1986 reserve schools for the first time were able to apply for funds to provide services to their students identified as having special needs. INAC currently distinguishes between low- and high-cost special needs students. Low-cost special education funding for reserve schools is provided by INAC as

a per capita grant and is not dependent on the identification of specific individuals or needs. Indeed, low-cost funding is intended to provide services such as resource teachers, remedial programs, and assessment services that may serve a broad range of the student population. As of 1993-1994 the base allocation was $274 per full-time student per year (plus adjustments for reserve isolation and size factors).

Students can only be identified as high-cost after they are formally assessed by a psychologist or other specialist and only after the results of this assessment, accompanied by an Individualized Education Plan (IEP), are forwarded to INAC for approval. Once approval has been received, students are identified on the Nominal Roll registration form for the following year. Since 1991-1992, INAC has not required schools to formally retest students for whom a second consecutive year of funding is being requested (1993-1994 funding levels for high-cost students are shown in Table 5).

Since the new funding regulations came into effect there has been a noticeable increase in the number of formally identified special needs students attending reserve schools. Between 1985 and 1988 there was an increase of more than 200% in the number of Indian students identified as high-cost special needs in Manitoba (Working Margins Consulting Group, 1989, see Table 3). In 1988-1989 7.1% of the total student enrollment was identified as special needs. In Canada the percentage of the student population having special education needs has been variously reported in different studies as 9.07% (Canadian Council for Exceptional Children, 1989), 9.7% (Alberta Education, 1977), 12.5% (Roberts & Lazure, 1970) and 15.5% (Council of Ministers of Education, 1983). The comparable percentage for Manitoba is thought to be approximately 10% (Manitoba Education and Training, 1989). The relatively low proportion of identified special needs students may be attributed to the lack of special education resources, both human and financial, that are available for Indian students attending schools on reserve, which is in turn a reflection of the level of priority attached to special education services.

The above statistics, together with the review of indicators of risk among Indian students, provide an indication of problems in the identification of

Table 3
Growth in High Cost Indian Special Education Enrollments
Manitoba, 1985/86 to 1988/89

Year	*High Cost Student Enrollment*	*Percent of Total Student Enrollment*
1985/86	322	2.2
1986/87	471	3.2
1987/88	786	5.2
1988/89	1,075	7.1
Change (1985 to 1988)	+234%	

Note. Data are for Indian students whose permanent residence is on reserve. Students attending all school types (band, federal, and provincial) are included.
(Source: Working Margins Consulting Group, 1989, p. 48).

special needs students in Indian schools. However, it is necessary to look at the actual programs provided in specific schools in order to gauge the degree to which special needs are being met. The following section provides the results of such an examination for five band schools in Manitoba.

Evaluation of Special Education Programs in Five Band-operated Schools

In 1987 the Manitoba region of INAC established a funding program through which reserve schools were to be evaluated every five years (Mariash, 1989). This section summarizes the findings of five schools evaluations that took place between 1989 and 1992.[3] (In order to protect the anonymity of the schools involved, the following summary does not provide specific information concerning the numbers of teachers and students in individual schools.)

All but one of the schools were located in northern Manitoba. Three of the schools were in communities that did not have year-round road access. The enrollments of the schools evaluated ranged from about 80 to about 600 students. Three of the schools provided programs from kindergarten through grade 12, one school terminated at grade 8, and one at grade 9. Collectively, the enrollment of the five schools was 1,481 at the times of the surveys. Total teaching staff in the five schools, including teacher aides and principals, numbered 110, of whom a total of 41 were interviewed, including all teachers who were directly concerned with special education classes. Several other nonteaching staff were also interviewed in each school as noted below, such as home-school coordinators, nurses, band councillors, and others. A total of 115 special education students were identified in the five schools, including 50 in grades K-3, 38 in grades 4-6, 18 in grades 7 and 8, and 7 in grades 9-12. All available folders for these 115 students were examined as part of the evaluation process.

The following questions guided the evaluation of the special education programs at each school.

1. How are students referred for special education in the school? Are appropriate assessment and placement procedures used? Are these procedures being properly and consistently implemented?
2. How are Individual Educational Plans (IEPs) devised for special needs students, and are these adequately followed up?
3. What consultative or specialist resources and materials are used by the school in support of the classroom teachers? Are these resources available and adequate to meet the needs that have been identified?
4. To what extent have special education policies been developed concerning assessment, placement, informed consent, monitoring and follow-up of IEPs etc.?
5. Does the school have appropriate resource and curriculum materials to support the special education program?
6. To what extent are parents consulted with or involved in the special education process?

The special education evaluations included:

a. Informal on-site interviews with various teaching and administrative staff.

b. An examination of existing special education/resource folders of special needs students, including psychological reports and Individual Education Plans (IEPs).
c. Informal Interviews with various resource personnel in the Department of Indian Affairs, the tribal council responsible for placing selected special needs students in private homes in Winnipeg, and selected former staff of the school (in cases where key school staff had recently left the school at the time of the evaluation).
d. A tour of the school's facilities and selected observation of resource room activities.

The observations obtained during the evaluation of the five band-operated schools are summarized below. It should be noted that the five schools were experiencing similar problems with respect to special education. There was some variation in the quality of the Individual Educational Plans and in the nature of special needs programming that was provided, but there was a consistent lack of monitoring, specific guidelines for teachers, and follow-up of the recommendations. The comments below are generally true of all schools evaluated except where otherwise noted.

1. The process of identifying and assessing students with special needs may be described as follows. School staff in band-operated schools (resource teacher, classroom teacher, principal) target selected students for special education in May. Parental approval for comprehensive assessment is secured and a psychologist is contracted and flown into the community to assess the targeted students. The psychologist completes his or her assessment (she or he may also make a recommendation for a further speech and language assessment), writes a detailed report for each student and makes a recommendation regarding funding required from INAC. Based on the psychologist's report, the resource teacher develops an IEP for each student. A Student Specific Request Form, the student's Psychological Assessment Report and the IEP are forwarded to INAC as a request for special needs funding for the following academic year.
2. Reasons for referral typically provided in the students' folders included severe academic delay and stuttering (suspected Tourette's Syndrome), severe articulation difficulties, suspected cerebral palsy, and violent/aggressive behavior. Tests commonly administered by the psychologist included the Brigance Diagnostic Inventory of Basic Skills, the Torrence Test of Creative Thinking, the Peabody Picture Vocabulary Tests, the Wechsler Intelligence Scale for Children—Revised, the Bruininks-Oseretsky Test of Motor Proficiency, and the Key Math Diagnostic Arithmetic Test.
3. When the IEPs of the currently identified special needs students were examined, very little evidence of monitoring student progress was found. Individuals responsible for developing and carrying out student goals and objectives were not always identified on the IEP forms. In particular, the role of the classroom teacher in supporting the program developed by the special needs/resource teacher was not always delineated. Beginning and completion dates for specialized intervention related to each objective were also sometimes missing. As a result, IEPs seemed to be used primarily as a

means for obtaining funds from INAC rather than serving to assist the child in classroom intervention.

4. IEPs were sometimes too skeletal in form. In some cases, IEPs were photocopied and used for a number of students with similar needs. Although the intent may have been for teachers to expand and personalize these skeletal IEPs, evidence for this was not consistently present.
5. Although parental approval is required prior to assessment, this procedure was not always followed. For example, parents' signatures on approval forms were sometimes dated after assessments had taken place.
6. There were no specific instructional guidelines for teachers to follow in the students' Psychological Assessment Reports. The main recommendation in each of these reports was that the student be funded by INAC, but other recommendations often fell short of practical programming directives. The few recommendations that were provided were often too broad in scope. For example, for a student who displayed a severe language delay, it was recommended that "X be encouraged to copy, write and draw as much as possible to develop language skills and a language base." For a student who displayed very low self-esteem, it was recommended that "the teacher reward 'little successes' in the classroom." For a student who was achieving substantially below his grade level, it was recommended that he "be provided with individualized programming materials in the mainstreamed classroom and participate in small group activities, with the assistance of an aide."
7. Special needs programming for the identified high-cost students varied but primarily took the form of remedial instruction or alternative programming. Sometimes it was provided by the resource teacher on a pull-out basis; in other cases it was provided by the classroom teacher and aide. It was found that:
    a. Teachers were sometimes unclear about their role regarding programming.
    b. Teachers were not always involved in or informed about the students' IEPs.
    c. Withdrawal times for students to go to the resource room were inconsistent.
    d. Regular consultation meetings between the classroom and resource teachers were not always scheduled into their timetables. A collaborative/cooperative resource model was not always in place.
    e. Maintaining any specialized programming for a student was often difficult. The resource teacher often expected the classroom teacher to take more initiative; the classroom teacher expected the resource teacher to be continuously involved in the delivery of programming for the special needs student.
    f. In some schools there were limited resource materials, particularly for the upper grades. If special materials were purchased, they were not properly catalogued, making them difficult for other teachers to access.
8. Few students were identified as having special needs in the senior high school grades (grades 9-12, see Table 4). From 8.0% to 11.9% of students in the primary, intermediate and junior high grades were identified as special

Table 4
Percentage of Students Identified as Special Needs by Grade Level in Four Reserve Schools in Manitoba
Various Years Between 1989 and 1992

	School				
Level (Grades)	*A*	*B*	*C*	*D*	*Weighted Average*
Primary (K-3)	7.5	12.9	4.2	16.3	8.0
Intermediate (4-6)	5.6	12.3	16.9	5.0	11.9
Junior High (7-8)	3.1	6.1	13.3	0.0	8.5
Senior High (9-12)	1.6	0.0	2.4	—	2.2
Total	5.5	11.1	7.4	10.8	7.6
Total Identified Special Needs Students			104		
Total Enrollment			1,366		

(Sources: Data from various school evaluation studies completed between 1989 and 1992).

needs students, while only 2.2% of those in the senior high grades were special needs students in the four schools for which complete data were available. This pattern could reflect a greater dropout rate among special needs students, thus leaving few such students in the high school grades.

9. There was an evident absence of specialist support for the educational program. For example, if psychologists were flown in to an isolated community to conduct assessments, their primary role was assessment, not consultation on programming, teaching strategies or curricular resources. Many teachers expressed a need for more guidance in these areas.
10. This concern was compounded by a lack of training of resource teachers. Although all resource teachers were certificated teachers, few had special education certificates.

Organizational Support of Special Education Services in Band Schools

Apart from the funding supports for low- and high-cost special education students described above, INAC provides no additional funds for special education consultants, clinicians, or administration. INAC does provide one consultant and one psychologist who provide assistance to reserves throughout the province in establishing their programs. Bands typically provide a community-centered special education program consisting of resource rooms and special classes, such as remedial classes, life skills classes, and reentry classes for former dropouts, allowing students to stay in their home communities while attending school. However, the resource room and special class teachers and the regular classroom teachers operate without special education clinician and program support (see Figure 5). Lacking a system of coordinators or clinicians and other program supports, band-operated schools use private consultants to provide special education support services and student assessments.

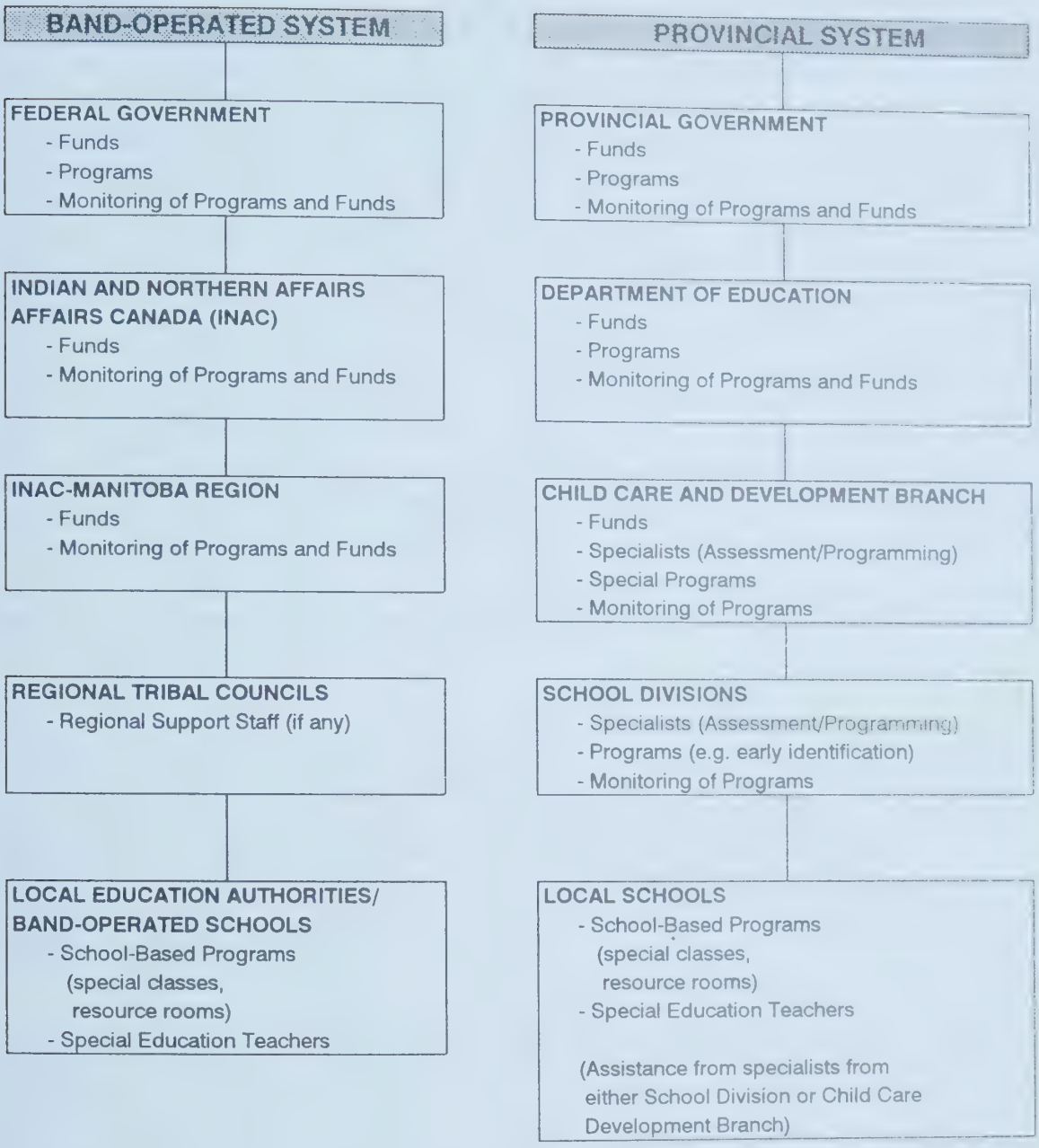


Figure 5. Comparison of band-operated and provincial special education support and monitoring systems.

Provincial (Manitoba) Special Education Policy

INAC support for special education, as with other aspects of the education program, is intended to meet provincial standards (Education and Research, 1985). Therefore, it may be useful to compare the INAC system of supports to the provincial system. Since 1967 the province of Manitoba has required school divisions or districts to provide educational programs for children with special needs. There is mandatory provincial legislation regarding special education (Canadian Council for Exceptional Children, 1992). The goal of special education in Manitoba is "to support children in developing the knowledge and skills they require to live meaningful, self-fulfilling lives with as much independence as possible in their communities" (Manitoba Education and Training, 1989, p. 3). This support is evidenced through a variety of special education programs and services to provincial school divisions/districts.

The province of Manitoba, through its department Manitoba Education and Training, provides special education funding to provincial school boards. The funding is provided on the basis of the total general student population as well as the number of individually identified special needs students (Canadian Council for Exceptional Children, 1992). Funding support for special education includes four components: coordinator and clinician support and Levels I, II, and III support (Manitoba Education and Training, 1992).

The coordinator and clinician support is based on the school division's total enrollment. This support is calculated as the "lesser of eligible enrollment divided by 700 adding one for any remainder or the number of qualified personnel employed, multiplied by $45,000" (Manitoba Education and Training, 1992, p. 9). In other words, divisions must spend this money on qualified personnel or they will lose it. Level I support is the lesser of the school division's enrollment divided by 180, multiplied by $45,000; or allowable expenditures. Level II support ($8,520 per student) is for individual students who have been identified as "severely multi-handicapped, severely psychotic or autistic or profoundly deaf" (Manitoba Education and Training, 1992, p. 9). Level III support ($18,960) is for individual students who have been identified as "profoundly multi-handicapped" (Manitoba Education and Training, 1992, p. 9).

Aside from special education support based on either total school enrollment or individually identified students, Manitoba Education and Training provides additional supports to provincial schools. These supports include direct consultation services for the visually and hearing impaired, vision/hearing screening, adaptive learning materials and special equipment, regionally based psychological and speech/language support services, the Diagnostic Centre (which provides assessment and programming), the Manitoba School for the Deaf, and professional development workshops.

In summary, the government of Manitoba, through Manitoba Education and Training, has developed a comprehensive system of special education

Table 5
Comparison of Provincial and Federal Funding Policies
(Prior to the 1994/95 Fiscal Year)

Type of Support	*Manitoba Education and Training*	*Indian and Northern Affairs (INAC)*
Base Funding	Enrollment/700 (rounded upward to nearest whole number) x $45,000	Enrollment x $274* (plus adjustments for isolation and size)
Level I Funding	Enrollment/180 x $45,000 (equivalent to $250/student)	$4,560/Identified Special Needs Student*
Level II Funding	$8,520/Identified Special Needs Student	$8,520/Identified Special Needs Student
Level III Funding	$18,960/Identified Special Needs Student	$18,960/Identified Special Needs Student

*As of the 1994/95 Fiscal Year, Base Funding and Level I Funding have been combined in a single grant of $400 per student. This is composed of $281 per student for Base Funding plus $119 per student for Level I Funding. The new funding policy also provides that no school will have its special education funding reduced as a result of the new formula.

Table 6
Special Education Low Cost and Level I Funding Available for Hypothetical Band Schools, Based on INAC Funding Policies (Prior to the 1994/95 Fiscal Year)

Band-Operated School Enrollment	*Level I Students*		*Low Cost Grant*	*Level I Funding*	*Total Funding Available*
	Percent	*Number*			
200	5%	10	54,800	45,600	100,400
200	10%	20	54,800	91,200	146,000
600	5%	30	164,400	136,800	301,200
600	10%	60	164,400	273,600	438,000

services (see Figure 5). A wide range of funding, specialist, and program supports is available for students with special needs attending provincial schools. Schools, whether large or small, are able to obtain specialist and program assistance from either a school division or from Manitoba Education and Training.

Comparison of Federal and Provincial Special Education Systems

Support services provided by the provincial government for special education are different from those provided by the federal government. These differences are found both in the amounts of funding provided to schools and school divisions, and in the degree to which the respective governments exercise control over the way the money is spent. Table 5 provides a comparison of types of funding support provided in the two systems in Manitoba during the study period (1989-1992).[4]

It is difficult to compare the level of funding provided under the two systems, except by showing how they might work in practice with given enrollments. A typical smaller band school might have an enrollment of about 200 students, whereas a larger school might have an enrollment of about 600. From Table 4 it would be expected that each band school might identify 5% to 10% of their students as requiring special support, and most of these students would be assessed as requiring Level I support. Given this, the schools might be provided the amount of funding for special education (leaving aside Level II and III support) as shown in Table 6.

The comparable types of funding available through the provincial system to a given school are the clinician/coordinator support and Level I support. Provincial clinician funding is based on the collective enrollment of the school division. Therefore, the amount of support that might be considered as the share of the individual school would be calculated as a proportion of the division's total funding based on the proportion of the total enrollment attending the school. In Table 7 hypothetical enrollments for divisions as well as schools are shown, and these enrollments have been selected to show the possible range of situations that might occur.

Table 6 suggests that the funding available for special education programming, apart from Level II and III funding, in band schools with an enrollment of 200 might be from $100,400 to $146,000, depending on the number of iden-

Table 7
Special Education Funding Available for Hypothetical Provincial Schools Based on Manitoba Education Funding Policies

Provincial School Enrollment	Divisional Enrollment	Clinician Funding		Level I Funding	Total Funding[a]
		Division	School		
200	200	45,000	45,000	50,000	95,000
200	800	90,000	22,500	50,000	72,500
600	600	45,000	45,000	150,000	195,000
600	1,500	135,000	54,000	150,000	204,000

[a]School share of funding only.

tified Level I students. As shown in Table 7, for provincial schools the individual school's share of funding, apart from Level II and III funding, would be in the range of about $72,500 to $95,000, depending on the size of the school division in relation to the size of the school.

For a band school with an enrollment of 600, special education funding available might be from $301,200 to $438,000 (see Table 6), whereas for provincial schools of this size it might be from $195,000 to $204,000 (see Table 7). These figures do not take into account additional amounts provided to band schools because of small size or geographic isolation.

In spite of similar or higher funding levels for special education at band schools in comparison to provincial schools, in practice the overall system of supports for special education is much stronger in the provincial school system. Provincial schools all have access to special education consultants, clinicians, or coordinators through their school divisions. Provincial schools can also obtain more specialized assistance from regionally and provincially based consultants employed by Manitoba Education and Training. Some additional services are provided directly by the province, such as the Diagnostic Centre and professional development workshops. In short, the combination of school level, divisional, and provincial supports for special education results in a comprehensive system that provincial schools can access.

In contrast to the provincial system, INAC funding guidelines do not specifically provide for such services as administrative support, back-up specialist support, and early identification programs, nor does INAC provide a regional support system (see Table 8).

The lack of coordinators and clinicians in the band-operated system results in:

1. funds being spent on assessments by outside private consultants;
2. an absence of ongoing program consultation to special education and regular classroom teachers; and
3. lower numbers of identified special education students.

Given the funding policies reviewed above it appears that Indian education authorities have access to the financial resources required to provide a more comprehensive special education program, at least in theory. The question,

Table 8
Comparison of Special Education Services Provided to Students Attending Provincial and Band-Operated Schools in Manitoba

Special Education Service	*Provided in Provincial Schools?*	*Provided in Band-operated Schools?*
Remedial Education	Yes	Yes
Student Assessment	Yes	Yes
High Incidence Funding	Yes	Yes
Low Incidence Funding	Yes	Yes
Itinerant Teachers	Yes	No
Administrative Support	Yes	No
Specialists for Assessment and Programming	Yes	No
Early Identification Programs	Yes	No
Regionally-based Support	Yes	No
Provincially-based Support	Yes	No

(Source: Tom Walker, Director of Education, Swan Lake Education Authority, personal communication, Nov. 12, 1990).

then, is why is there evidently not an equivalent system in support of band schools? At least five factors seem to be involved.

Band schools are small and tend to work in isolation. Half of the schools located on reserves in Manitoba have enrollments of fewer than 200 students, and more than 70% have enrollments of fewer than 300 (Unpublished INAC Nominal Roll data, 1992). Moreover, these small schools are scattered throughout the rural and northern areas of the province. Although there is a system of tribal councils to which most Indian bands belong, the tribal councils have little responsibility in the area of elementary-secondary education. Only one tribal council in Manitoba has organized a school division (the Southeast Tribal Division for Schools) to provide central administrative and educational supports for a group of four reserve schools.

Local knowledge of special education, and local planning skills generally, are weak. As was found in the evaluation of five special education programs, there is often a need for professional development in the area of special education. Small schools may often lack the expertise to develop a special education program or even to organize systematic assessment of students. In order to both obtain and effectively utilize the special education funds available to them, band-operated schools would need to systematically undertake testing of students and would have to carefully allocate the per capita grant to either hire a specialist to do such testing or contract for such testing to be done. Once special needs students are identified and funding is received, this funding could be used to further enhance the special education program and services in the school. In effect, this is a two-stage funding process that involves delays and uncertainty as the school awaits funding approval by INAC. This process presents a barrier for some schools wishing to access funding.

Unlike provincial school divisions, band-run schools are not required to employ clinicians or to meet established program standards. As noted above, provincial

funding policy requires school divisions to employ qualified special education personnel in order to receive special education funds earmarked for the purpose. Bands, which are not subject to such a requirement, may use special education money for other purposes at the expense of the special education program. At present no system monitors whether this is happening.

Neither INAC, the various tribal councils, nor the chiefs have established any regional or province-wide education service organizations in support of band schools. Individual bands could allocate part of their education funding to establish regional support organizations that might be equivalent to the provincial system of supports. INAC could, as well, reserve part of the special education funds available for the support of such a regional organization. However, this has not happened, and it is not clear that bands or Indian education authorities would agree to such an arrangement as they would have to relinquish part of their local educational budgets.

Although funding for identified Level I students in band schools appears to be available, there may be a cap on this funding regardless of demonstrated need. Some school administrators believe that if they exceed the expected identification rate of 10%, they might not be able to access additional funds from INAC regardless of the demonstrated need.

Conclusions and Recommendations for Improving the Delivery of Special Education in Band-operated Schools

The foregoing discussion leads to four general conclusions. First, there is evidence that the need of Indian students for special education services may be greater and more complex than the needs of the general student population in Canada. Evidence of greater needs among Indian students takes the form of higher rates of age-grade deceleration, higher withdrawal rates, lower scores on achievement tests, and less favorable socioeconomic circumstances among Indian students in comparison with others. The greater educational and social needs of Indian students are complicated by language and cultural differences that may or may not be appropriately taken into account by the schools. There is little evidence that special education services in particular are provided in a way that has taken the social and cultural context of Indian students into account.

Second, Indian communities and schools are ambivalent in their attitudes toward special education programs. There is much concern about negative labeling and streaming of Indian students. Parents, educators, and Indian leaders are wary of remedial programs that tend to limit Indian students' opportunities and are concerned about the disproportionate number of Indian students who have been placed into such programs in provincial schools. They are also critical of standardized tests that are based on Canadian norms and are believed to be culturally biased. At the same time, those working in Indian schools are keenly aware of a variety of special needs among their students and wish to have the best, most suitable program to meet these needs. Moreover, identification of special needs students is a way of obtaining additional resources for the school.

Third, there are serious problems in the delivery of special education programs in Indian schools. The problems with the delivery of special education programs in the five schools evaluated were remarkably similar. These

problems involve the definition of the special education program, the roles and responsibilities of staff, the direction and monitoring of psychologists' work, the monitoring and followup of students, and the involvement of parents in the process. These problems may be related in part to the ambivalent attitude mentioned above, which tends to make special education programming a low priority in the school. In turn, the lack of development of special education programs means that special education programs in Indian schools are often based on outdated remedial program models, which may reinforce the community's negative view of the value of special education programs.

Fourth, there is no regional support system to assist local Indian schools as they attempt to identify special needs students and develop programs for them. Currently there is no system of monitoring the implementation of special education programs, apart from infrequent overall school evaluations. The lack of regional support or monitoring again undermines the local program and tends to reinforce negative perceptions of special education. It also tends to isolate each school system, and with many Indian schools having very small enrollments, there is little room in the budgets of these schools to hire specialists or address special needs.

In summary, local control of Indian education has not solved the problems of special education program delivery in reserve schools and has only recently begun to address them. The energy and enthusiasm that sometimes accompanies the movement to local control can neither be expected to solve these problems, nor to continue over a longer time frame. Local control has provided an opportunity to introduce new programs and concepts that may be a better reflection of local needs and the local cultural context. However, for the most part, at least in the realm of special education, the locally controlled schools do not seem to have taken advantage of this opportunity.

In some ways, local control has increased the difficulty of implementing new or specialized programs by isolating communities to a greater extent than they were under federal administration. It is unrealistic to expect small, isolated school systems to undertake educational program development without some type of support or assistance from outside the community. Moreover, locally controlled educational systems are more subject to local political pressures than those operated on a regional or province-wide basis. Given the special education evaluations summarized above, it is evident that not all the money earmarked for special education purposes was used for these purposes in some of the schools evaluated. Indeed, Indian education funds in general are not always protected from the demands of local band councils that they be used for other purposes (Hall, 1992).

What, then, should be done to provide more effective special education services in band-run schools? Based on the findings, we suggest the following five interrelated steps.

Develop a regional service network to support local programs. At present there does not appear to be an appropriate network of support services to assist individual schools, which are scattered in rural and northern areas throughout the province, to develop local policies, programs, and services. These support services must be able to provide specialized assessment, program development, and training assistance to local schools that require it. The provincial

system of support services in Manitoba provides an example of the kinds of needs that are to be met.

In addition, as shown in this article, Indian schools' special education programs are in need of monitoring and evaluation. Although a broad school evaluation once every five years is important, it is equally important that individual programs, including special education, receive regular feedback on their operation and performance. Because there is often a lack of local expertise in this area it would be useful to have a regional service that could provide such monitoring to those schools that require it.

A regional support network would also provide for sharing of information, program ideas, and services. Cooperation among communities is difficult under local control; a regional network would provide an avenue for cooperation.

Ensure that special education program funding fully addresses Indian needs. This review identifies some concerns over the funding of special education on reserves during the study period (1989-1992). First, although there seemed to be adequate provision for local special education services when the per capita grant and Level I funding were combined, schools seemed to have difficulty in fully accessing Level I funding. This may have contributed to the lack of development of various special education programs and services.

INAC has recently adopted a new per capita funding formula intended to cover both base funding and Level I funding. This new system should streamline the budgeting process and lend itself to improved local program planning. The new funding formula seems to be roughly equivalent to the provincial funding level, but it also seems to be a reduction in total funding available.

The local delivery system for special education needs to be developed. There are currently many inadequacies in the delivery of special education services in reserve schools. Although a regional support network would be helpful, the responsibility for developing these services rests with local Indian education authorities. In this respect there are three key factors:

1. Principals and directors of education must provide leadership in communicating the need for and purpose of special education in the school.
2. An educational program for staff and parents must take place, with the assistance of outside expertise where needed.
3. Special education funding must be protected from demands from other areas of educational programming or from politically based demands to use education funding for other purposes.

Provide resource teachers, and teaching staff generally, with inservice training and skill upgrading related to special education. Many resource teachers are not fully trained and experienced, and many other teaching staff are not clear about the role of the resource program and the way in which IEPs are developed and implemented jointly by classroom and resource teachers. Inservice training in this area is clearly needed, and in some cases resource teachers may need to be encouraged to enroll in more extensive professional upgrading.

Establish clear program standards and guidelines for special education in band-operated schools. Program standards or guidelines are needed to ensure that students in need of special education services have access to such services on reserves. These standards also need to be monitored, and the results of the

monitoring process should at least be reported to the community. There are three ways in which the establishment and monitoring of special education program standards could take place.

1. A province-wide Indian educational body could be formed by interested Indian education authorities to establish these standards and undertake the responsibility of monitoring them;
2. INAC could make special education funding conditional on spending the funds in specific ways, and back this up by requiring specific monitoring and reporting activities; or
3. INAC and/or a province-wide Indian educational body could contract with Manitoba Education and Training to monitor compliance with guidelines, which would be identical to provincial guidelines.

We note that the Educational Framework Agreement between the Assembly of Manitoba Chiefs and the government of Canada seeks to establish a province-wide educational body by 1995. This may provide the opportunity to create an effective system of supports and standards in the area of special education for reserve school systems in Manitoba.

Notes

1. The term *Indian* is used here to refer to the status or registered Indian population, whereas the term *Aboriginal* is meant to be inclusive of status and non-status Indians, Metis, and Inuit.
2. Manitoba withdrawal rates among Indian students may be higher than those of Indian students in other parts of Canada. Hull (1987, p. 44) found that there were higher rates of school leavers among Indian students in the western provinces than in central or eastern Canada, with Manitoba's student leaver rate being third highest among Canadian provinces.
3. These evaluations were conducted by Eleoussa Polyzoi, one of the co-authors of this article. Dr. Polyzoi's specific role was to examine and evaluate the special education programs at these schools. On average she was on site for approximately four to five days for each evaluation as a member of a team.
4. As of the 1994-1994 fiscal year INAC has instituted a new funding formula that provides a $400 per capita grant taking the place of the per capita grant base funding plus Level I funding. In other words, individual schools are not required to undertake an assessment process to justify Level I funding. The $400 is composed of $281 for base funding (a 2.5% increase over the 1993-1994 funding level) plus $119 for Level I funding. If 10% of the student population needed Level I support, the cost under the old funding system would have been $456 per student. If only 5% of students needed Level I support, the cost would have been $228 per student. The new funding formula, therefore, appears to be a substantial reduction in overall funding for special education.

References

Alberta Education. (1977). *The Alberta special education study: Condensed report*. Edmonton, AB: Alberta Education, Planning and Research.

Assembly of First Nations. (1988). *Tradition and education: Towards a vision of our future* (vol. 1). Ottawa, ON: Author.

Assembly of Manitoba Chiefs. (1984). *Our tomorrows today: Different paths to quality Indian education in Manitoba*. Winnipeg, MB: Author.

Assembly of Manitoba Chiefs. (1991). *Indian education funding study* (Final Report). Winnipeg, MB: Author.

Canadian Council for Exceptional Children. (1989, October). *Keeping in Touch: A Quarterly Newsletter from the Canadian CEC Office*. North Bay, ON: Nippissing University College.

Canadian Council for Exceptional Children. (1992, December). Provincial/territorial survey of services to exceptional children. *Keeping in Touch: A Quarterly Newsletter from the Canadian CEC Office*.

Council of Ministers of Education. (1983). *Survey of special education in Canada 1982-1983*. Toronto, ON: Author.

Education and Research. (1985). *A study team report to the task force on program review*. Ottawa: Supply and Services Canada.

Hall, D.R. (1992). FED-BOS: The federally controlled band operated school and the no-policy policy. *Canadian Journal of Native Education, 19*, 67-66.

Hull, J. (1987a). *An overview of the educational characteristics of registered Indians in Canada*. Ottawa, ON: Indian and Northern Affairs Canada.

Hull, J. (1987b). *An overview of Registered Indian conditions in Manitoba*. Ottawa, ON: Indian and Northern Affairs Canada.

Hull, J. (1990). Socioeconomic status and Native education in Canada. *Canadian Journal of Native Education, 17*(1), 1-14.

Indian and Northern Affairs Canada (INAC). (1987). *Classification and housing economic categorization of Indian bands by zone*. Ottawa, ON: Author.

Indian and Northern Affairs Canada (INAC). (1992a). *Basic departmental data—1992*. Ottawa, ON: Author.

Indian and Northern Affairs Canada (INAC). (1992b). *Indian register population by sex and residence, 1991*. Ottawa, ON: Author.

Island Lake Tribal Council. (1987, March). *Report on special education in the Island Lake Region*. Winnipeg, MB: Author.

Koens, P.R. (1989). Local control within a model of the sociology of knowledge: A study in curriculum. *Canadian Journal of Native Education, 16*(1), 37-44.

Manitoba Education and Training. (1989). *Special education in Manitoba: Policy and procedural guidelines for the education of students with special needs in the public school system*. Winnipeg, MB: Author.

Manitoba Education and Training. (1992). *Funding of schools: 1992/93 school year*. Winnipeg, MB: Author.

Mariash, L. (1989). *School evaluations: The effective schools cycle*. Winnipeg, MB: Indian and Northern Affairs Canada, School Evaluations/Education Planning.

McBride, C., Gagne, M., & Atwell, K. (1990). *Characteristics of public administration employment on-reserve; 1986 Census*. Ottawa, ON: Indian and Northern Affairs Canada.

McKnight, B. (July 14, 1988). *Letter from Bill McKnight, Minister of Indian Affairs and Northern Development, to all Indian bands and councils with attached Policy for the Provision of Education Facilities and Outline of the School Space Accommodation Standards*. Unpublished document.

Morris, A. (1862-1979). *The treaties of Canada with the Indians of Manitoba and the North-West Territories*. Toronto, ON: Coles.

National Indian Brotherhood. (1972). *Indian control of Indian education* (Policy paper presented to the Minister of Indian Affairs and Northern Development). Ottawa, ON: Author.

Paquette, J. (1986). *Aboriginal self-government and education in Canada* (Aboriginal Peoples and Constitutional Reform, Background Paper Number 10). Kingston, ON: Queens University, Institute of Intergovernmental Relations.

Phillips, R.S., & Cranwell, C.R. (1988). The need for community-based special education programs in band-operated schools of Manitoba. In A.A. Ortiz & B.A. Ramirez (Eds.), *Schools and the culturally diverse exceptional student: Promising practices and future directions* (pp. 112-119). Reston, VA: Council for Exceptional Children.

Roberts, C.A., & M.D. Lazure. (1970). *One million children: A national study of Canadian children with emotional and learning disorders*. Toronto, ON: Crainford.

Stevens, H. (Principal Investigator). (1982). *A review of changes in the living conditions of the registered Indian population of Manitoba during the 1970's*. Winnipeg, MB: Social Planning Council of Winnipeg.

Working Margins Consulting Group. (1989). *Evaluation of Indian counselling services in Manitoba: Final report*. Unpublished report prepared for a steering committee of Manitoba tribal councils and Indian and Northern Affairs, Manitoba Region.

The Alberta Journal of Educational Research *Vol. XLI, No. 1, March 1995, 63-83*

Grace V. Malicky

and

Charles A. Norman
University of Alberta

Perceptions of Literacy and Adult Literacy Programs

There are numerous definitions of literacy in the literature but relatively few data on the views of adult learners in literacy programs regarding literacy and even fewer on the views of adult literacy instructors. This study examined the views of 94 learners and 31 teachers regarding literacy as well as the actual classroom experiences of a subgroup of learners. Results showed that learners tended to view literacy from a fundamental perspective but entered programs for job-related reasons. Their teachers viewed literacy from a functional perspective but presented programs that were fundamental in nature. Emancipatory views were reflected to a very limited extent in either views of literacy or actual classroom experiences.

Bien qu'il y ait de nombreuses définitions de l'alphabétisation citées dans plusieurs documentations, il existe peu de données des opinions qu'ont les apprenant(e)s adultes participant aux programmes d'alphabétisation en ce qui concerne l'alphabétisation comme telle. Il existe encore moins de données sur les opinions qu'ont ces gens de leurs instructeurs et instructrices. Cette étude examine les opinions de 94 apprenant(e)s et 31 instructeurs et instructrices en ce qui concerne l'alphabétisation ainsi que les expériences en classe d'un sous-groupe d'apprenant(e)s. Les résultats indiquent que les apprenant(e)s considèrent l'alphabétisation d'une perspective fondamentale mais s'inscrivaient dans ces programmes pour des raisons reliées au marché du travail. De leurs côtés, les instructeurs et instructrices percevaient l'alphabétisation d'une perspective plutôt fonctionnelle mais présentaient des programmes de façon fondamentale en sorte. Des opinions émancipées ont été exprimées d'une façon très limitée lorsqu'il s'agissait soit des expériences en salle de classe, soit des points de vue concernant l'alphabétisation.

Introduction

Although some information is available on the concepts of reading held by adults in literacy programs (Gambrell & Heathington, 1981; Norman & Malicky, 1986), little research has been done on how instructors in adult literacy programs view literacy or on the congruence of views of literacy learners and teachers. In light of dropout rates from adult literacy programs that have been estimated to be as high as 70% (Quigley, 1992), it is important to understand not only the degree of consistency between learners and instructors in how they view literacy, but also the degree of consistency of their views with

---

Grace Malicky is a professor in the Department of Elementary Education and Director of the Reading and Language Center. Areas of research interest include emergent literacy, clinical reading, and adult literacy.
Charles Norman is a professor in the Department of Educational Psychology. Areas of research interest include emergent literacy, cognitive development, and adult literacy.

literacy practices in classrooms. Street (1992) stresses the importance of understanding the nature of the relationship between ideas and practices. He believes that literacy is a social, political phenomenon that always involves power relations.

There is considerable support in the literature for a sociopolitical view of literacy (Fingeret, 1992; Harman, 1987; Scribner, 1984). Kazemek (1988), for example, describes literacy as a personal and social process of coming to know that is political, cultural, and context-dependent. He decries current adult literacy programs that employ teacher-controlled, didactic methods and argues instead that "literacy is an ethical endeavor that has as its goal the liberation of people for intelligent, meaningful, and humane action upon the world" (p. 467). He would support a position where one first determines the goals of literacy education and then designs a program consistent with these goals.

There is often considerable difference in the backgrounds of adults in literacy programs and of their teachers. In general, learners tend to come from more disadvantaged backgrounds than their teachers (Hunter & Harman, 1979). In a study of the literacy practices in three communities in Piedmont, Carolina, Heath (1983) found that the programs and practices used in schools were often at variance with the cultural backgrounds and values in people's lives. When teachers modified programs to make them more consistent with the language and literacy backgrounds of the children in their classes, achievement levels of children from disadvantaged backgrounds increased.

A major purpose of the study reported in this article was to examine the perceptions of both learners and teachers regarding the nature and goals of literacy as well as their views regarding the literacy programs in which they were involved as teachers and learners. The focus of this part of the study was on the degree of congruence between the views of learners and teachers. Lack of agreement in perceptions of literacy and literacy programs could be one factor in high dropout rates from adult literacy programs. A second purpose of the study was to observe the nature of the experiences of a subgroup of learners in literacy classes to determine the extent to which programs reflected the goals and views of learners and teachers. A discrepancy between stated goals and the means to realize them could negatively affect the success of literacy programs. Analysis of data was conducted from the perspective of the following three major concepts of literacy that are widely reflected in the literature and have dominated the work of UNESCO from 1945 to the present (Lind & Johnston, 1986).

Concepts of Literacy

In 1946 the concept of fundamental education was adopted by UNESCO with the core content of adult education programs involving skills of thinking, speaking, listening, calculating, reading, and writing. Illiteracy was viewed as a disease to be eradicated and the major tool was primary school education. Adult literacy education was viewed as an extension of schooling and often the same materials and techniques were used. Few writers define literacy this narrowly today, although fundamental literacy is sometimes included as one aspect. For example, Valentine (1986) uses the term *general* literacy to refer to the reading and writing aspects of literacy and contrasts this with *functional* literacy. As Kazemek (1988) points out, this view of literacy is still reflected in

the instructional practices of many literacy programs. For the purposes of this study, fundamental literacy refers to the skills of reading and writing.

The concept of functional literacy began to influence international thinking in 1956 with Gray's survey of reading and writing conducted for UNESCO. Whereas Gray stressed the relative nature of literacy and the need to relate literacy training to the context in which individuals live, UNESCO linked the concept of functional literacy with social and economic development. The major focus was on literacy for work and increased productivity. Literacy was not viewed as an end in itself, but rather as a way of preparing people for social, civic, and economic roles. Literacy was viewed from a human capital paradigm of economics in which education is viewed both as a form of consumption and a form of investment. The goal of education from this perspective was to enhance human capital, which would in turn increase productivity and yield a positive rate of return. Although this narrow view of functional literacy is still reflected in some programs, many writers view functional literacy from a much broader perspective. For example, Valentine (1986) defines it as "an individual's reading and writing ability in relation to the reading and writing tasks imposed by, or existing in, the environment in which that individual resides and seeks to function" (p. 109). Levine (1982) focuses more on information when he defines functional literacy as "the possession of, or access to, the competencies and information required to accomplish those transactions entailing reading and writing in which an individual wishes—or is compelled—to engage" (pp. 263-264). In this study functional literacy is viewed from the broader perspective of Valentine and Levine, although it is important to note that one of contexts in which people deal with written language involves work.

Although functional literacy continues to dominate much of the literature in North America, UNESCO began to move away from a functional definition in 1975. The Declaration of Persepolis adopted at the International Symposium for Literacy in 1975 retained the relative and contextual aspects of literacy but stressed the political, human and cultural aspects as well. Literacy was defined as

> not just the process of learning the skills of reading, writing and arithmetic ... Literacy creates the conditions for the acquisition of a critical consciousness of the contradictions of society in which man [sic] lives and of its aims; it also stimulates initiative and his participation in the creation of projects capable of acting upon the world, of transforming it. (Hamadache & Martin, 1986, pp. 128-129)

This is the definition used in this study for the terms liberation or emancipatory literacy. From this perspective, literacy is not viewed as neutral, "for the act of revealing social reality in order to transform it, or of concealing it in order to preserve it, is political (Hamadache & Martin, 1986, p. 128). This liberation or emancipatory view of literacy reflects the work of Freire (1970) and others (Hunter, 1982; Kretovics, 1985; Stanage, 1986) who have focused on empowerment and the ability to bring about change in inequities and injustices in society.

Generally, the goal of literacy programs based on functional and fundamental views of literacy has been to adapt individuals to fit into the existing

sociopolitical order and hence to preserve this order. The goal of liberation or emancipatory literacy, on the other hand, is for individuals to act on society and bring about change.

Procedures

The Sample and Context of the Study

The sample of learners consisted of 94 adults enrolled in literacy programs in one urban center. The gender, age, status, and mean reading level of subjects are presented in Table 1 along with the type of literacy programs in which subjects were enrolled. The range of reading grade scores on the Tests of Adult Basic Education (TABE, 1987) was from level 2.1 to 9.2. The higher proportion of females to males was reflective of the population in adult basic education programs in the area as was the relatively high proportion of immigrants.

A subgroup of 18 learners was randomly selected for more in-depth study, and the teachers of these learners were also involved in the research. Half the subjects in this subsample were immigrants and the other half Canadian born. Twelve were female and six were male. Eight of these subjects remained in adult basic education, bridge (a transition between ABE and high school classes for ESL students) or high school classes across the duration of the study, whereas the others either went to different types of programs or left for personal reasons (e.g., health, family, etc.). Of the five who went to other programs, four went into trades or pretrades programs whereas the other went to a high school program in another setting because of dissatisfaction with his previous setting.

Most instructors of adult basic education, bridge, and high school programs who taught the 18 subjects in the subsample were interviewed to obtain their views of literacy and to describe typical literacy classes. Views on literacy were

Table 1
The Sample of Literacy Learners

Variable	*Number or Mean*
Gender	
Females	61
Males	33
Mean Age	29.5
Status	
Canadian-born	40
Immigrants	54
Attendance in the Literacy Program	
Full-time	69
Part-time	25
Nature of the Literacy Program	
Formal	84
Non-formal	25
Mean Reading Level on TABE	5.0

obtained from 26 instructors, 19 of whom were teachers in adult basic education or bridge classes and seven of whom were in high school programs. Five additional instructors provided descriptions of typical literacy classes.

Data Collection and Analysis

Data were collected from the sample of 94 adults attending literacy programs and from 31 teachers of these programs. Both formal testing and interviews were used to collect information from learners. The Test of Adult Basic Education (TABE) was used as a screening tool to select learners below a high school reading level because the major focus of the study was on adult basic education. All learners were interviewed at the beginning of the study using a structured interview schedule to obtain demographic data as well as information about their concepts of literacy and reasons for entering the literacy program. This interview schedule was based on a questionnaire used by Davis and O'Brien (1985) with adults in literacy programs in Nova Scotia. The purpose of this interview was to describe the adult participants in literacy programs as well as to identify the meanings they associated with literacy. The major changes to the interview schedule involved omitting some questions because Davis and O'Brien had noted that it took too long to administer. The initial interviews were audiotaped and responses were recorded on the interview sheets.

Across the three-year time period of the study learners were also interviewed at six-month intervals while they attended literacy programs and at the time they exited literacy programs (if they did). One focus of these interviews was on what they viewed as helpful in the literacy program and what else they would like to have seen in the programs. Two open-ended questions were used to elicit these comments: What do you think are the strengths of what you are doing in your classes? and What else would you like to do that you are not now doing in your classes? Although change across time was not of major concern in relation to these questions, the three-year time period of the study was essential to understand the relationship of participation in literacy programs with employment (Malicky & Norman, 1994b) and to provide data on more than one class for most learners in the study. Research assistants conducted interviews under the supervision of one of the principal researchers and all interviews were audiotaped and transcribed for analysis.

Learners in the subgroup of 18 were observed approximately every three months in their classes to obtain data on the nature of their actual classroom experiences. These observations were conducted for as long as the participants remained in literacy or high school upgrading classes across the three-year time frame of the study. Fieldnotes were kept of these observations, outlining what materials were used as well as providing a running commentary on what students and teachers did and said during the lessons.

The teachers of the subgroup of 18 were interviewed to determine their concepts of literacy as well as their perceptions of the literacy program they were providing. Questions that focused on ways in which literacy and literacy programs were perceived were selected from a questionnaire for literacy instructors developed by Davis and O'Brien (1985). Examples are: Why do you think being literate is important? What do you think constitutes being a literate person? What do you think should be the purpose of literacy programs like the

one in which you are involved? Would you describe a typical class or session for me? All interviews were audiotaped and transcribed.

Data collected from the total sample of 94 learners on the TABE and initial interview were quantified and tabulated where possible. Means were calculated to provide a description of the adults in the study. Responses to other questions such as literacy concepts and perceived strengths of programs were categorized and the number of responses falling into each category determined. Data gathered on follow-up interviews with learners, on interviews with instructors, and through classroom observations were read to identify emerging categories regarding perceptions of literacy programs by program participants, views of literacy held by instructors, content of programs and the nature of interactions in classrooms.

Learners' Perceptions of Literacy and Literacy Programs

The majority of learners were full-time students at the start of the study, attending classes during the daytime five days per week. Part-time students generally attended classes once or twice each week. Most of the adults in the study were from formal programs, with 51 attending literacy classes in a vocational center and the other 33 attending classes in a continuing education program sponsored by a school system. The remaining 10 adults were in nonformal literacy programs, five in a volunteer one-to-one program and the other five in community-based programs.

Funding for the programs involved in this study was provided by two government departments—Manpower and Education—with the result that these programs generally reflected one of the following two orientations: (a) a manpower orientation concerned primarily with enabling individuals to participate in the labor force, and (b) a social demand orientation concerned with enabling individuals to participate more fully in all aspects of society. In this study the literacy program situated in the vocational center had an explicit manpower orientation whereas the other programs had more of a social demand orientation.

Definitions of Literacy and Goals for Entering Literacy Programs

Program participants were asked directly to provide their definition of literacy during the initial interview (Are you familiar with the word literacy? What do you think it means?). Many were unfamiliar with the term, particularly those with English as a second language. Of the 36 who did respond, 25 or 69% included reference to reading, writing and/or literature in their definitions. Five referred to knowledge, five to education, schooling or learning, and one to English as a second language. Generally, definitions provided by participants tended to reflect a fundamental notion of literacy.

Further information on participants' views of literacy was obtained by asking them, Why did you enter this program? Many participants gave several reasons, but the most common were job-related: 83% believed that increased literacy could help them improve their job opportunities. In addition to job-related goals, several people cited personal/psychological reasons such as feeling better about themselves and developing self-confidence. These reasons were given more frequently by Canadian-born respondents (43%) than by immigrants (17%). Social reasons, such as meeting people or becoming more

independent, were given by approximately one quarter of the men and immigrant women, but 73% of Canadian-born women provided reasons in this category. This may reflect differences in men's and women's ways of knowing with women perceiving themselves more in terms of connections and relationships (Kazemek, 1988). Both general and specific educational goals were also important to the participants in this study. Immigrants (43%) gave more specific goals, such as learning how to read and improving English skills, than did Canadian-born respondents (16%).

Perceived Strengths of Literacy Programs and Suggested Changes

Participants were asked at regular intervals to comment on their perceptions of the programs they were attending. Some participants left the program before they could be given a follow-up interview, and hence the data presented in Table 2 were collected from 83 of the sample of 94 learners. Because some participants moved from one program to another during the study, comments on both basic literacy and high school upgrading programs are categorized in Table 2. The scores in the table indicate the number of comments received regarding areas that participants felt were strengths or in need of change at some point during their participation in programs.

Areas of Language Arts

The information presented in Table 2 indicates that participants frequently referred to traditional areas of language arts education as both strengths of programs and areas in which more instruction was needed, although this was less true for Canadian-born males than for the other program participants. The areas that participants perceived as most helpful were reading and writing, with 39% citing writing as a strength of their programs and 28% citing reading. The other two areas mentioned frequently were grammar (24%) and vocabulary (20%). It is not surprising that these two areas were viewed as valuable by immigrants because most were learning English as a second language, but it is interesting that several Canadian-born women (23%) included these areas as well. One woman talked about improving her language competence because she had a young child, and it may be that some of the other women also felt the need to present better language models for their children. In any case, this finding is consistent with the relatively high number of social goals given by

Table 2
Perceptions of Literacy and Upgrading Programs: By Learners

Focus of Comment	*Number of Comments Received*	
	Strength	*Area of Suggested Change*
Language arts areas	121	57
Specific instructional materials	5	0
Content	6	9
Specific activities	7	6
Interpersonal aspects	33	15
Organization of classes	48	26
Level/Amount/Pacing	3	26

Canadian-born women for entering literacy programs. None of the Canadian-born men cited vocabulary or grammar as strengths of or needs in their programs.

When asked what changes they wanted in their programs, women tended to provide more responses than did men, but even the women gave relatively few responses to this question. In relation to areas of language arts, two areas mentioned by several of the immigrant females involved more focus on speaking (22%) and listening (16%), again consistent with expectations for ESL learners. These areas were mentioned by only two of the Canadian-born participants. More reading, writing, spelling, and pronunciation were also mentioned by a small number of both immigrants and Canadian-born participants.

The relatively heavy focus on traditional areas of language arts education is consistent with definitions of literacy provided by program participants and confirms that a fundamental or traditional view of literacy was widely held. Hence it appears that participants brought this view with them to their literacy programs, although persistence of the view could also reflect the nature of the programs in which they were enrolled.

Content, Materials, and Instructional Activities

In general, skill areas such as reading and writing were mentioned much more frequently than content or topics. Only six participants identified general or specific content as a strength of their programs (e.g., the environment, Meech Lake, different life styles) and six noted general or specific content that they would like to have had included (e.g., books on animals, news, history of different nations). Three other participants cited problems with their programs related to content: two talked about content being useless or irrelevant and one identified a specific workshop on how to dress as a waste of time. Overall, though, content was not a strong focus of the program participants in this study, and there was little difference between immigrants and Canadian-born participants or between males and females in the extent to which they identified content or topics as strengths or needed changes in their programs.

Similarly, specific instructional materials or activities were rarely mentioned as either strengths or weaknesses of programs. A few respondents referred to specific books being used in their classes as good and to activities such as story projects and research papers as useful. They offered no suggestions regarding what materials they would like to use and tended to identify activities they didn't like rather than make more constructive suggestions. This is perhaps not surprising in that they would not have been aware of the range of options available.

Interpersonal Dimensions

Instead, learners frequently focused on interpersonal dimensions when talking about strengths and areas of needed change. Both immigrants (41%) and Canadian-born participants (53%) indicated that their teacher was a real strength of their program. Many simply indicated that their teachers were good, but others talked about what made their teachers good, for example,

> The teacher, I thought, was an excellent teacher. She knew what she was talking about and she wouldn't go any further unless everyone in the class understood it. (female, Canadian-born)

> I like the teachers ... they are real helpful and friendly and willing to help every time I ask them. (female, immigrant)

> All the teachers are really good. They really pushed me further than I ever thought I would go.... When I was at school before, they treated me like a child and I didn't get any benefit from it. Now they treat me like an adult and you know, you can talk to teachers like, when I was at school you didn't know the teacher's first name. (male, immigrant)

The somewhat higher percentage of Canadian-born than immigrant participants who cited their teacher as a strength of their program might reflect the heavier focus of Canadian-born respondents on personal/psychological goals for entering literacy programs. Interpersonal relationships would appear to be more critical to meeting these goals than would specific content or instructional activities.

It is important to note that not all comments about teachers were positive. Seven students felt that they needed more time or assistance from their teachers, and three others felt they were being treated like children. According to one woman,

> She teaches as if we're in an elementary class type of thing.... Like she, you know, it's like "Okay kids we're going to do this, this and this today." Like, I find that writing spelling words out is a real pain. I mean that's what grade ones and twoers do type of thing. (female, Canadian-born)

Two of the immigrants felt they were being treated unfairly by their teachers, for example,

> Sometime I think it maybe—prejudiced. I think the teachers ... I copy from the encyclopedia word by word, and they give me the paragraph back full of red marks. (female, immigrant)

Relatively few comments related to interpersonal relationships with other students. Some students such as one Canadian-born male talked about how he enjoyed "going to school and talking with the students," but generally the participants in this study viewed their teachers as being key to their literacy learning.

Organization of Classes and Pacing of Instruction

Program participants also commented on the way programs were structured, particularly those students in the computer-based literacy program. Initially, many were enthusiastic about working on computers (11 commented on this as a strength of their programs), but once the programs were underway, five people expressed concern that they were not really learning to use computers, for example,

> I really thought the computer program was gonna be like where you learn the computer, like doing it so you go out and get a job. (female, immigrant)

A need for class instruction along with use of computers was identified by three individuals and six indicated that they would like more class time. Typical of these comments were the following:

> It's fun working on the computers but if you wanted to sit down there and learn, I find you learn more on an individual classroom basis. (male, Canadian-born)

> I like the classroom teaching you know. In the classroom, I mean, there are more things, like a teacher teaching you and additional things that one might add and you get more knowledge from that. (female, immigrant)

Not only those participants in the computer-based literacy program expressed a need for more class time. Five participants expressed a desire to access pull-out classes for extra assistance.

Several comments were related to level, pacing, and amount of material covered. Two participants were pleased that they were challenged and one that the material was easy. However, several were concerned that the program moved too slowly (6), that the material was too simple (8), or that there was too much review (7). Twelve of these comments were received from immigrants, several of whom had attained more advanced levels of schooling in their countries of origin, but some Canadian-born participants also felt programs did not move quickly enough for them, for example,

> We've worked for weeks on these prepositional phrases and subordinate clauses and all these new words. And she's sort of drilled that into us and now it's becoming quite boring. If she doesn't move to some new material soon, I don't know how much more the class will put up with it. (female, Canadian-born)

Far fewer participants felt that programs moved too quickly (2), presented too much (2) or didn't provide enough review (1).

The Subsample

The subsample of 18 subjects was involved in more in-depth study, including interviews with their teachers and observations in their classrooms. The major purpose of this part of the study was to obtain an indication of the views of literacy held by literacy instructors and the nature of literacy programs provided. It was also conducted to determine the degree of consistency between teachers' and learners' beliefs and between what teachers said about literacy and what they actually did in their classrooms.

Eleven of the students in the subsample were attending literacy classes in the vocational center at the beginning of the study. Most of these students were immigrants and most had taken ESL courses at the center prior to entering adult basic education classes. Of those who were in basic literacy classes at the beginning of the study, two began in a level 3-4 classroom, one in level 4-5, three in level 6-7, and two in level 7-8. Two other students were in bridge classes for the first observation. Nearly all these students were observed in more than one level, and five eventually moved on to high school classes. There were strict attendance requirements in this program and movement from one level to the next was based on test performance.

Six of the students attended literacy classes in the school system-sponsored adult basic education program. Students were involved in both computer-based and classroom instruction. Two of the students moved into the high school program sponsored by the same school system and one went on to a pretrades program at the vocational center. The final student began her program in a community-based literacy program and then entered the high school program at the vocational center.

Instructors' Views of Literacy and Literate People

As indicated above, 26 instructors of the 18 adult learners in the subsample were interviewed to find out how they viewed literacy. They were asked to tell why they thought literacy was important, what they thought constituted being a literate person, and what they thought the purpose should be of literacy programs like the one in which they were involved. Most instructors provided multidimensional descriptions of literacy and of literate people.

The most common descriptions given by instructors involved functional aspects of literacy; a total of 44 responses fell into this category (several informants gave more than one functional aspect of literacy). Typical responses in this category follow:

> To make people more employable.
>
> Handle forms and most things you run into in most situations.
>
> Can read and write well enough to function in society.
>
> Function in daily activities.
>
> Being able to get along or succeed in day-to-day living.

Fundamental descriptions of literacy were also prevalent in the data; 27 of the responses fell into this category. Most involved a reference to a specific area or aspect of language arts, but several also involved references to schooling or academic achievement.

> Basic understanding of spelling, punctuation and basic paragraphs.
>
> Basic abilities for reading and understanding the written word.
>
> Literature, reading and writing.
>
> Preparation for [other programs named].
>
> Reading is not just decoding but enriching what a person already knows.
>
> Writing as a means of communicating.

A similar number of responses focused on the sociopolitical (11) as the personal, psychological (11) aspects of literacy. Examples of sociopolitical responses included: "if you're not literate, you're third class," "the illiterate let somebody else take control of their lives," "to make people who are not easily governed." Psychological aspects of literacy included: "able to enjoy themselves," "gives a sense of self-esteem," and "self-confidence." Five respondents recognized that literacy is relative, depending on both the person and context. One, for example, said "it depends on what the person wants to do." Finally, eight responses reflected a more global/general view of literacy, such as "want a future," "aspire to a better life," "important in every facet of life," and "related to being human." When views of literacy were examined in relation to whether teachers taught in adult basic education, bridge, or high school contexts, there was little difference in how literacy was viewed.

What is most striking is that all instructors interviewed for the study included some aspect of functional literacy in their descriptions of literacy or literate persons. Most also included a fundamental notion of literacy in their descriptions whereas less than one quarter referred to emancipatory or psycho-

logical outcomes of literacy. One would anticipate on the basis of these results that most literacy programs would reflect a somewhat equal emphasis on functional and fundamental aspects of literacy, whereas there would be far less of a focus on sociopolitical aspects or personal development.

Two types of information were collected to determine the degree of correspondence between views of literacy espoused by instructors and the programs they provided for students. First, they were asked to describe typical class lessons, and second, approximately every three months a lesson was observed and fieldnotes taken.

Descriptions of Typical Lessons

Instructors were asked to describe a typical class and the following probes were used if necessary: What are the books and materials you use? How much time do you spend on each area? Descriptions of typical lessons were obtained from 31 instructors (nine instructors provided more than one lesson description because they taught two or more different courses or levels). Nineteen instructors taught adult basic education classes (two of these also taught high school classes), four taught bridge classes (one of these also taught high school classes), and eight taught high school classes. Adult basic education and bridge classes in the vocational center were frequently taught by teams of two instructors.

Most instructors included more than one area in their descriptions of typical classes, and these were categorized as shown in Table 3. In a level 7-8 class at the vocational center, for example, the teacher indicated that the students worked on reading in the morning and writing in the afternoon. Reading instruction was based on a workbook containing factual articles; students completed questions for homework, and these were corrected and discussed in class. Literature was also included in the program, with students reading stories and answering questions in class. Writing instruction was focused around two workbooks, one with an emphasis on grammar and the other on vocabulary.

In the adult basic education program sponsored by the school board, students spent a considerable portion of their time in computer-based literacy programs. In the Pathfinders program, students completed pretests and on the basis of their results the computer provided them with a list of materials to use. After completing the recommended exercises the students took a posttest on the computer. Teachers were available to answer questions both on the use of the computers and on what students were doing. Classes varied considerably across instructors; reading and writing were common components, with spelling, grammar, oral presentations, listening, and study skills included in various classes.

In the community-based literacy program the teacher included one half-hour every day on reading comprehension and also what she referred to as exercises on an area of need identified from the students' writing, for example, quotation marks. In addition, students worked for one and a half hours Mondays on writing, Wednesdays on reading comprehension, and Fridays on spelling. The teacher selected the materials for the students to read, and current topics served as the basis for brainstorming and writing.

Table 3
Aspects Included in Descriptions of Typical Classes

Aspects of language/literacy	*Percentage (number) of Instructors*					
	ABE (N=24)		*Bridge (N=7)*		*High School (N=21)*	
Grammar	33%	(8)	29%	(2)	29%	(6)
Vocabulary	19%	(4)				
Oral presentations	13%	(3)	14%	(1)	5%	(1)
Discussions	19%	(4)				
Reading						
Strategies	21%	(5)			10%	(2)
Literature	19%	(4)	14%	(1)	57%	(12)
Individual	21%	(5)	29%	(2)		
Question/Answer on reading	33%	(8)	43%	(3)		
Writing						
Skills	13%	(3)	43%	(3)		
Paragraphs/essays	58%	(14)	29%	(2)	33%	(7)
Journals	13%	(3)				
Study Skills	8%	(2)	14%	(1)		
Other	13%	(3)				

From the data in Table 3 and these few examples, it is apparent that in spite of the heavy emphasis on functional literacy evident in the definitions provided by instructors, descriptions of typical adult basic education classes focused almost exclusively on fundamental aspects. This was true to a large extent for high school classes as well, although there were some differences. The majority of lessons at the high school level included work on literature and writing, whereas those in adult basic education classes involved a wider range of aspects of language arts. Typical of programs provided at the high school level is the following description provided by a teacher in the vocational center setting. The program she offered included five major components: short stories, novels, plays, research projects, and eight book reports. The program was planned to meet the requirements of the provincial curriculum.

Although it is possible that instructors were aware of the discrepancy between their definitions of literacy and the programs they provided, only a small number of them commented on this inconsistency during interviews. One, for example, indicated that the goal of the program should be to increase employability and self-esteem but "instead we are preparing students for an academic program." Another expressed concern that some students

> leave here feeling like a failure because they haven't made progress academically. They still aren't any better off in terms of a job.... We may be doing a disservice by making people feel worse than they did when they came in.

She noted that "we are changing" but that students were not always in favor of this change.

> Sometimes when we give students the opportunity to talk about what will help them in their everyday lives, that's not what they want from us. They don't see education as providing that.

Both of these instructors wanted a heavier focus on functional aspects of literacy. A third instructor felt that the institution was focusing on both functional and fundamental aspects and she wanted to see a heavier emphasis on sociopolitical goals.

Classroom Observations

Data obtained through actual classroom observations were also analyzed to gain further insight into the degree of congruence between descriptions of literacy provided by instructors and what they actually did in their classrooms. These data were analyzed both in relation to the content of the lessons as well as the nature of interactions between teachers and students. The number of lessons observed in adult basic education settings was 42, the number in bridge settings seven and the number of high school classes 21.

Content

Table 4 provides an indication of what the content focus was in the lessons observed. Again, most of what was observed in classrooms could easily be classified according to traditional areas of language arts education. Students were frequently involved in reading and writing activities, although the nature of these activities varied depending on whether they were attending adult basic education, bridge, or high school classes. Students generally read a wider range of materials in adult basic education classrooms than in high school English classes where the focus was primarily on literature. Writing activities were similar across the two contexts, with the exception of journal writing, which was not observed in high school classrooms. Very few bridge classes were observed, so it is difficult to draw any generalizations about what was happening in those classes. It does appear, however, that there was a heavy focus on grammar in these classes probably because of the heavy enrollment of ESL students. In some of the bridge classes observed students spent the entire time working through a grammar book, completing exercises and correcting them together in class.

Teachers generally selected what students would read and assigned topics for writing (although for writing teachers sometimes gave several possible topics for students to choose from and left the possibility open for students to choose one of their own instead). Journal writing and independent reading opportunities did provide for some degree of student control over content, but these were far more common activities in adult basic education than in high school classrooms.

What Table 4 masks is that functional or emancipatory goals may have been served in some of the reading and writing activities. Short stories occasionally lead to heated discussions about topics such as child abuse or drunk drivers. Suggested writing topics were also sometimes controversial in nature and designed to heighten consciousness about issues, for example, "Women should/should not have equal job opportunities." Generally, however, teachers reported little input from students into curriculum decisions and most

Table 4
Aspects of Lessons Observed in Various Contexts

Aspects of language/literacy	*Percentage (number) of Lessons*		
	ABE (N=42)	*Bridge (N=7)*	*High School (N=21)*
Grammar	29% (12)	100% (7)	19% (4)
Vocabulary	12% (5)	29% (2)	
Listening	7% (3)		
Oral presentations			5% (1)
Pronunciation	2% (1)	29% (2)	
Reading			
SRA Kits/Independent Reading	5% (2)	14% (1)	5% (1)
Literature	19% (8)	14% (1)	50% (10)
Newspaper	19% (8)		
Specific Skills	17% (7)		14% (3)
Writing			
Journals	12% (5)		
Paragraph/essay	12% (5)	14% (1)	29% (6)
Specific Genre	5% (2)		5% (1)
Specific skills	12% (5)		33% (7)
Life/Study skills	5% (2)		
Subject areas	10% (4)	29% (2)	

discussions were related to skill development rather than to issues in the life worlds of the students in the classes.

In general, there was considerable correspondence between what teachers said they did in typical classes and what actually went on in classrooms; activities in classrooms were primarily focused on fundamental aspects of literacy. Bridge teachers demonstrated a heavier emphasis on grammar than they reported, but there were few other striking differences between what teachers reported and what they did.

Interactions

Classroom observations were also analyzed in relation to the nature of the interactions that occurred between students and teachers, and the results of this analysis are presented in Table 5. The types of interactions placed in the "other" category included such things as the teacher giving students words and asking them to find corresponding pictures, teachers writing questions on the blackboard and having students answer them orally, and the teacher reading words and having students repeat them. What Table 5 fails to capture is the sequence of interactions in classrooms. In order to provide some indication of this, observations of selected classes for two of the students are briefly described.

Sandra was observed first in a community-based literacy program and then moved to a high school program in the vocational center. At the beginning of the first class observed in the community-based literacy program (March), students were working at circular tables on reading comprehension exercises

Table 5
Interactions Apparent in Classroom Observations

Nature of Interaction	Percentage (number) of lessons ABE (N=42)		Bridge (N=7)		High School (N=21)	
Teacher lectures	19%	(8)	43%	(3)	33%	(7)
Teacher questions, students answer (oral)	33%	(14)			24%	(5)
Teacher reads, questions/students answer	29%	(12)			14%	(3)
Teacher/students do example together	19%	(8)	14%	(1)	29%	(6)
Teacher goes through test/assignment orally	14%	(6)			33%	(7)
Students do exercise independently, check orally in class	26%	(11)	43%	(3)	29%	(6)
Oral discussion	24%	(10)	29%	(2)	10%	(2)
Students read orally	5%	(2)	14%	(1)		
Students do exercise in groups	12%	(5)	29%	(2)	29%	(6)
Students work individually on assigned task	50%	(21)	14%	(1)	24%	(5)
Students do test	7%	(3)	29%	(2)	5%	(1)
Students work individually on task of own choice	2%	(1)			14%	(3)
Other	14%	(6)	14%	(1)		

while the teacher worked with one individual on a possessive pronoun exercise. The teacher then handed out an exercise on quotations. She put an example on the board and directed the students' attention to a page in the handout, asking them to tell what was wrong with the sentences. The students chimed in answers as the teacher read each sentence. On the next page she read the directions and asked the students to do three items individually. She then called on individual students to provide answers. The teacher moved to reading comprehension and explained that she wanted the students to generate a main idea statement for articles on garbage and then go through the 5W questions. She solicited the 5Ws from the class and wrote them on the board. She then gave a different newspaper article to each of three groups, checking with each group as they worked and commenting on the appropriateness of the statements they had written. Each group then shared the main idea they had written with the rest of the class, going through the 5Ws after they had done so.

The next fall Sandra was in a grade 10 English class in the vocational center. The tables were organized in rows facing the front of the room and a text entitled *College Writing Skills* was used during the lesson observed. The teacher asked students to turn to a particular exercise in the textbook and to do it. They were to rewrite a passage that had nine errors (run-ons, dangling modifiers, etc.) in it. The teacher circulated helping individual students as the others worked quietly on their own. When the students finished, they passed their papers to another student, and the teacher went through the sentences one by one, asking students if there were any mistakes. After checking to determine how many students got all the sentences right, the teacher assigned the next

exercise on subject-verb agreement in the text. The students completed it individually, passed the papers to another student and again marked it together. The teacher then said, "Let's give you a chance to talk for awhile." He asked the students to write two topics on two pieces of paper. Students pulled topics from a box and spoke in front of the class for 30 seconds. The teacher encouraged the speakers by saying, "Well done," "Good job," and the students rated each other on a 10-point scale.

The second student, Mary, began upgrading in the program sponsored by the school system and then enrolled in English 30 in the vocational center. During one class in the upgrading program (March) the teacher began by working through an exercise on prefixes and their meanings from a workbook entitled *Communication Skills*. The teacher then assigned an exercise on suffixes and the students completed it, helping each other figure out the directions and then working quietly. Mary frequently consulted the dictionary as she completed this assignment. The teacher went through the exercise with the students who chimed in with answers. He then gave them the following writing assignment on the board:

> Take one of the words you have worked with in lessons 5-8 and expand it to give it more meaning, that is understanding of it to your reader.
>
> OR
>
> Take one of your own.

The teacher explained the assignment and that the students would be editing their own writing. The students wrote until the end of class time.

From data in Table 4 and from these examples, it is evident that most classes revolved around activities selected by the teachers that students did and then frequently checked orally with the teacher. Most interactions among people were teacher-directed although students did do some teacher-assigned activities in groups and in four of the classes observed had an opportunity to work on a task of their own choice. Overall, however, teachers tended to define what counted as knowledge and had most of the power in classrooms. There were some exceptions as indicated in the following description of Mary in a high school English class at the vocational center one year after the class described above.

The class began by the teacher helping students fill in a form on short stories they had read. The form contained the headings title, author, genre, and plot. One student mentioned that Solchuk means "son of sun" and the teacher said, "See I told you guys. You teach me something new everyday." Another student asked whether Mr. Solchuk was the protagonist or antagonist. Rather than providing an answer, the teacher said, "That's a good question. What do you guys think?" After considerable discussion the class came up with a definition. Similarly, other questions were dealt with by the class rather than by the teacher. When discussing one of the stories, Mary commented, "You want to believe parents do things like this." She talked about conflict in her family with her mother and father and indicated that she hadn't mentioned this in class because she didn't know that the young people would understand. A male student responded, "I think we can empathize with you," and discussion ensued about lack of family support. Finally, the teacher brought the discussion back to the story by asking the class to think for tomorrow about the fact

that the father of the character in the story was coming around. As the students packed up to leave, another student came over to talk to Mary and said he thought they had had similar experiences. Mary later told the researcher that she never used to talk in class, but this teacher had done something different with this class—they worked in groups and talked about things.

What this teacher had done was give up some of the ownership of knowledge to the students in her classroom. Her definition of literacy was quite similar to that of many other teachers interviewed—she gave fundamental (reading, writing, listening, speaking, viewing), functional (dealing with everyday life, being able to use resources, where to get help), and personal, psychological (increased self-confidence) goals for literacy. However, in her description of a typical class, she referred to the transactive theory of Rosenblatt (1978) and indicated that there was a heavy focus on response to literature. Students were encouraged to talk and write about how they felt about the literature they read, bringing in their personal experiences. This was not the only class in which a teacher relinquished some of the control of content and classroom interactions to students, but it was rare rather than common among the upgrading classes observed during this study.

Conclusions and Discussion

Overall, the results of this study revealed considerable discrepancy within and between student and teacher groups in how literacy was viewed and practiced. When the learners entered adult literacy programs, they tended to hold a fundamental view of literacy, probably as a result of their past school experiences. However, they persisted in this view as they attended literacy classes in spite of the fact that when asked why they had entered literacy programs they frequently cited job-related (functional) goals. This would appear to reflect a mismatch between the perceptions of literacy held by adults in literacy programs and their goals for entering these programs. However, it is unlikely that the participants viewed this as a mismatch. Instead, they appeared to believe that a traditional education was the avenue to improved job opportunities in spite of the fact that there is little evidence in the literature to indicate that increased literacy will have a positive impact on employment (Graff, 1987; Levine, 1986). Indeed, most of the learners in this study returned to the same low paying, temporary jobs that they had held before entering literacy programs (Malicky & Norman, 1994b).

In contrast to program participants, literacy instructors tended to view literacy from both functional and fundamental perspectives, yet presented programs that were primarily fundamental in nature. Hence the programs they were providing were more consistent with the views of learners regarding the nature of literacy than with their own views. This seems to reflect a gap between theoretical orientation and practice in the classroom. Some instructors recognized the discrepancy between what they believed should be done and what they were doing, but generally felt constrained by their institutions and, according to one instructor, the students as well. Despite somewhat different mandates of the vocational, school system-sponsored and community-based programs, literacy content and instruction was markedly similar. The degree of consistency in the type of literacy instruction across the different contexts appeared to reflect the pervasiveness and impact of the fundamental view of

literacy instruction and learning in the province. To some extent this was determined in high school upgrading classes by the mandated provincial curriculum, but there is no mandated curriculum for adult basic education classes. In light of high dropout rates in the study (Malicky & Norman, 1994a), the strong emphasis on fundamental literacy needs to be critically examined in relation to goals and purposes of programs.

Emancipatory views of literacy were reflected to a very limited extent in the views of literacy instructors and students or in the content and interactions in classrooms. Classrooms tended to be teacher-dominated, with teachers determining the content of instruction and evaluating student performance. Students generally expressed positive comments regarding their teachers and although they valued opportunities to interact with other students they worked individually rather than cooperatively. Overall, there were few opportunities for students to share power with their teachers, and this did not appear to be an area of concern for most of them. The purpose of the literacy programs in this study could be viewed as helping students to adapt to or fit into the existing sociopolitical order, and hence reproducing this order (Apple, 1982; Giroux, 1983), rather than equipping them to act as change agents on society. It is important to note that there were classes in which students engaged in the kind of consciousness raising dialogue that Freire (1970) believes is an essential element of change. Hence it was possible to reach emancipatory goals even in high school upgrading classes with a mandated curriculum.

There were relatively few differences between females and males or between immigrants and Canadian-born program participants in their views of literacy or their comments regarding literacy programs. Immigrants tended to focus somewhat more heavily on the traditional areas and skills of language arts whereas Canadian-born subjects were more concerned with interpersonal dimensions of literacy instruction, but a relatively large proportion of comments from all groups fell into these two broad categories. There were the expected differences of immigrants valuing work on grammar, vocabulary, speaking, and listening, but even here Canadian-born females also identified vocabulary and grammar as significant areas of instruction. There were few differences based on gender or immigrant status on how content, materials, specific activities, classroom organization, and pacing of instruction were viewed.

A significant finding of the study involved the importance attributed by adult literacy students to their teachers as compared with materials or methods. Teachers spend a considerable amount of time searching for the "best" series or instructional technique, and publishers try to fill this need. The adults in this study viewed their teachers as a significant strength of their literacy programs and valued the assistance and explanations they received during classroom instruction. In contrast, they rarely referred to specific materials or techniques when describing strengths of their programs. Hence the results of this study do not support a heavy emphasis on methodology or technology. This was particularly evident in students' responses to the computer-based literacy program. Although some were enthusiastic about the program throughout the study, many expressed a desire for more interaction with teachers after the initial phase of their involvement in the program. It appears,

then, that more emphasis needs to be placed on teacher-learner interactions in planning and implementing literacy programs rather than on searching for the "best" methods or materials. In particular, the results do not support replacing teachers with computer-based instruction.

Although the finding that students viewed their teachers as crucial and generally "good" will be of comfort to literacy instructors, it raises a question regarding the independence of adult literacy learners. One of the basic principles of adult education is to lead adult learners in the direction of becoming independent and self-directed, and yet the learners in this study appeared to be quite dependent on their teachers, viewing them as necessary to learning. In a study of adult illiterates reentering the learning context, McDermott (1982) found that self-direction and independence increased slowly as adults experienced success and developed confidence in themselves as learners. Thistlethwaite (1983) also identified lack of confidence in ability to learn as an obstacle to self-direction and independence. Many of the students in the present study were lacking in confidence at entry to literacy programs (Malicky & Norman, 1994a). This does not mean that we should treat adults like children; rather it means that adults will require assistance to overcome their fears and anxieties and to develop confidence in their ability to learn. Adult literacy students require supportive, successful learning experiences to become independent, self-directed learners.

It is important to keep in mind when interpreting the results of this study that they are more a reflection of societal views of literacy and literacy teaching than of individual teachers. The teachers in this study appeared to be no more empowered than their students; they were constrained both by societal views of literacy as reflected in their students and by the institutions in which they worked. The institutions in turn were constrained by the views of policy makers and funders of adult literacy programs. There is a need for major partners in the adult literacy enterprise—learners, teachers, administrators, policy makers, funders—to critically examine their views of literacy and literacy learning. Only in this way will programs move beyond the current almost exclusive focus on fundamental literacy to achieve some of the emancipatory potential of literacy learning. Only in this way will literacy programs help people to make rather than take their place in society.

Acknowledgment

This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada.

References

Apple, M.W. (1982). *Education and power*. Boston, MA: Ark Paperbacks.

Davis, N., & O'Brien, M. (1985). *Literacy programs and their participants in the Halifax-Metro area*. Halifax, NS: Mount Saint Vincent University.

Fingeret, H.A. (1992). Literacy in the USA: The present issues. In UNESCO Institute for Education (Ed.), *The future of literacy and the literacy of the future* (pp. 15-23). Hamburg, Germany: Editor.

Freire, P. (1970). *Pedagogy of the oppressed*. New York: Herder and Herder.

Gambrell, L.B., & Heathington, B.S. (1981). Adult disabled reader's metacognitive awareness about reading tasks and strategies. *Journal of Reading Behavior, 13*(3), 215-222.

Giroux, H.A. (1983). *Theory and resistance in education*. South Hadley, MA: Bergin & Garvey.

Graff, H.J. (1987). *The labyrinths of literacy: Reflections on literacy past and present*. London: Falmer Press.

Gray, W.S. (1956). *Maturity in reading: Its nature and appraisal.* Chicago, IL: University of Chicago Press.

Hamadache, A., & Martin, D. (1986). *Theory and practice of literacy work.* UNESCO.

Harman, D. (1987). *Illiteracy: A national dilemma.* New York: Cambridge Book Company.

Heath, S.B. (1983). *Ways with words: Language, life and work in communities and classrooms.* Cambridge: Cambridge University Press.

Hunter, C. (1982). Literacy for empowerment and social change. *Visible Language, 16*(2), 137-150.

Hunter, C.S., & Harman, D. (1979). *Adult illiteracy in the United States.* New York: McGraw-Hill.

Kazemek, F.E. (1988). Necessary changes: Professional involvement in adult literacy programs. *Harvard Educational Review, 58,* 464-487.

Kretovics, J.R. (1985). Critical literacy: Challenging the assumptions of mainstream educational theory. *Journal of Education, 167*(2), 50-62.

Levine, K. (1982). Functional literacy: Fond illusions and false economies. *Harvard Educational Review, 52,* 249-266.

Levine, K. (1986). *The social context of literacy.* London: Routledge & Kegan Paul.

Lind, A., & Johnston, A. (1986). *Adult literacy in the third world: A review of objectives and strategies.* Stockholm: Institute of International Education.

Malicky, G.V., & Norman, C.A. (1994a). Participation patterns in adult literacy programs. *Adult Basic Education, 4*(3), 144-156.

Malicky, G.V., & Norman, C.A. (1994b). Participation in adult literacy programs and employment. *Journal of Reading, 38,* 122-127.

McDermott, P. (1982). Conversational asides: The social context of an adult literacy class. *Journal of Reading Behavior, 14,* 461-474.

Norman, C., & Malicky, G. (1986). Literacy as a social phenomenon: Implications for instruction. *Lifelong Learning 9*(7), 12-15.

Quigley, A. (1992). Looking back in anger: The influences of schooling on illiterate adults. *Journal of Education, 174*(1), 104-121.

Rosenblatt, L.M. (1978). *The reader, the text, the poem: The transactional theory of the literary work.* Carbondale, IL: Southern Illinois University Press.

Scribner, S. (1984). Literacy in three metaphors. *American Journal of Education, 93,* 6-21.

Stanage, S.M. (1986). "Unrestraining" liberty: Adult education and the empowerment of persons. *Adult Education Quarterly, 36*(3), 123-129.

Street, B. (1992). Sociocultural dimensions of literacy: Literacy in an international context. In UNESCO Institute for Education (Ed.), *The future of literacy and the literacy of the future* (pp. 39-51). Hamburg, Germany: Author.

Tests of Adult Basic Education. (1987). Monterey, CA: CTB McGraw Hill.

Thistlethwaite, L. (1983). Teaching reading to the ABE student who cannot read. *Lifelong Learning, 7,* 5-7, 28.

Valentine, T. (1986). Adult functional literacy as a goal of instruction. *Adult Education Quarterly, 36*(2), 108-113.

The Alberta Journal of Educational Research *Vol. XLI, No. 1, March 1995, 84-102*

Richard Conte

Jac J.W. Andrews

Melanie Loomer

and

Gillian Hutton
University of Calgary

A Classroom-based Social Skills Intervention for Children with Learning Disabilities

A social skills intervention consisting of coaching, role playing, and information sharing was implemented in a classroom for children who attended a special school for learning disabled children. Twelve children (10 males and two females, average age 12.2 years) participated in the experimental program that was implemented over a six-month period by a clinical psychologist who collaborated with the classroom teachers. Fifteen students (9 males and 6 females, average age 13.3 years) participated in the control group in which no formal social skills program was implemented. Relative to controls, children who participated in the experimental program showed superior sociometric status and improved social problem solving skills in interview situations involving responses to provocations. In addition, for one of these situations there was a significant relationship between change in social problem solving and change in sociometric data. The findings are discussed in terms of the determinants of the effects of social skills interventions.

Un programme d'intervention en habiletés sociales qui consistait à diriger la pratique de ces habiletés entre les élèves, à encourager les jeux de rôles, et à l'échange d'information fut implanté dans une salle de classe pour élèves ayant des difficultés d'apprentissage. Douze enfants (10 garçons et deux filles, âge moyen 12,2 ans) ont participé au programme expérimental qui a été implanté sur une période de six mois par un psychologue clinique en collaboration avec des enseignant(e)s en salle de classe. Quinze élèves (neuf garçons et six filles, âge moyen 13,3 ans) ont participé au groupe contrôle dans lequel aucune formation d'habiletés sociales formelle ne fut introduite. Relatif au contrôle, les enfants qui ont participé à ce programme expérimental ont démontré un statut supérieur sur l'échelle sociométrique ainsi qu'une amélioration dans leurs habiletés à résoudre des problèmes sociaux en situation d'interview. Dans ces interviews, on demandait des réponses fictives à des situations provo-

---

Richard Conte is in private practice and is adjunct associate professor of clinical psychology. His interests are in the area of social skills, attention deficit disorder, program evaluation, and parenting.
Jac Andrews is an associate professor in Educational Psychology. He is interested in cognitive strategy-based assessment and instruction, social skills intervention, issues in the academic and social development of children with mild disabilities, and the teaching-learning interaction in secondary and postsecondary education.
Melanie Loomer is a clinical psychologist at the Learning Centre. She is involved in assessment and counseling with children and adults with learning and attention problems.
Gillian Hutton is a research assistant at the Learning Centre and has interests in the areas of social functioning and cognitive strategy instruction.

catrices. En plus, pour une de ces situations il y avait un rapport significatif entre le changement dans l'habileté de résoudre des problèmes sociaux et le changement des données sociométriques. Les résultats sont discutés en termes de déterminants des effets des interventions des habiletés sociales.

Increasingly, researchers have become interested in links between learning disabilities and social skills difficulties (Baum, Duffelmeyer, & Geelan, 1988; Dudley-Marling & Edmiaston, 1985; Gresham, 1988; Gresham & Reschley, 1986; Wiener, 1987). This trend is reflected by the inclusion of social skills deficits in definitions of learning disabilities that have been proposed both by the Interagency Committee on Learning Disabilities (ICLD, 1987) and the Association for Children and Adults with Learning Disabilities (ACLD, 1985). Although these definitions have not obtained widespread acceptance and the status of social skills as a learning disability remains a controversial issue (Conte & Andrews, 1993; Forness & Kavale, 1991; Gresham & Elliott, 1989), there is little doubt that the social problems of children with learning disabilities need to be addressed (Andrews, 1989; Bryan, 1991; Schumaker & Hazel, 1984; Vaughn, 1985; Vaughn & LaGreca, 1988).

As detailed in the recent review of McIntosh, Vaughn, and Zaragoze (1991), social skills research with children who have learning disabilities has tended to focus on interventions outside of the classroom setting. In these pull-out studies children with social difficulties are targeted for intervention and then withdrawn from the classroom for the purpose of social skills training. Although such studies have been moderately successful in improving children's social behavior (Schneider, 1992), there has been a notable lack of success in transferring social skills from the training environment to the classroom setting (Gorney-Krupshaw, Atwater, Powell, & Morris, 1981; Schumaker & Ellis, 1982; Whang, Fawcett, & Matthews, 1984). The lack of transfer in previous pull-out studies has prompted us to implement a social skills intervention in classrooms and to provide training to the teachers so that they can cope more effectively with the social needs of their students.

Although there is an increasing trend to provide educational services to learning disabled children in regular classes, it is the case that many children, particularly those who are severely affected, receive their education in full-time learning disability placements. One of the potential disadvantages of full-time learning disability placement is that the students may be exposed to negative modeling of social behaviors that could conceivably lead to a deterioration in social behavior over time. In support of this view, Bierman, Miller, and Stabb (1987) found a decline in positive peer interactions over time in socially rejected boys in a special education setting. Bierman et al. (1987) suggested that over time rejected children may actually show a decline in the overall level of peer acceptance as a function of familiarity with each other. One purpose of the present study was to determine if a social skills program could improve overall peer status in an intact classroom of children with learning disabilities.

Even in special schools for students with learning disabilities, there are different ways to provide social skills training. On the one hand it is possible to identify individual students with social skills difficulties and to remove them from the classroom for the purpose of social skills training, which is the approach most commonly used in previous research (see McIntosh et al., 1991). In

contrast, in the present study we have opted to provide social skills training in the school classroom and to expose all children in the classroom to the same social skills program. There are several potential advantages to this in-class approach. Teaching social skills in the classroom increases the likelihood that awareness of appropriate social skills will become part of the fabric of the classroom. Appropriate social skills can be reinforced throughout the day, and not just during the one or two hours a week that children attend social skills classes. An in-class social skills program may also make it possible to provide the kind of extensive treatment that is needed in order to have impact on the social difficulties of many children with learning disabilities. In this regard McIntosh et al. (1991) highlighted the brevity of the intervention period as a factor that has limited the effectiveness of social skills interventions. By conducting an in-class social skills program, it should be possible to embed the intervention in the school curriculum.

Our approach in designing the present program was to utilize features of programs that had been successful in previous pull-out social skills interventions. Specifically, two aspects of previous programs appear to be particularly salient: the use of a combined coaching/role playing intervention and the use of information sharing techniques between peers. First, with regard to coaching and role playing, research on social isolates (Oden & Asher, 1977) and aggressive children (Bierman & Furman, 1984) suggests that a coaching/role playing model of intervention can lead to relatively long-term gains in social functioning. Oden and Asher (1977) describe three components in their coaching and role playing methodology: verbal explanation of social skills, opportunities to practice the skills by playing with a peer, and a postplay review session. In their work with social isolates, they found that this coaching procedure led to long-term improvement in peer sociometric status. Gresham and Nagle (1980), Ladd (1981), and Ladd and Mize (1983) have also shown that coaching/role playing intervention significantly increased the number of positive peer ratings received by rejected children.

In developing the information sharing method that was used as a component of the present program, Bierman (1986) and Bierman and Furman (1984) have stressed the importance of having students share information about themselves and about their feelings with their peers. Fox (1989) showed that having students with learning disabilities share information about each other resulted in greater improvements in peer acceptance than a program in which students exchanged information about academic subjects.

In addition to the coaching/role playing and information sharing, some components of the program were directed specifically at two areas of deficiency in children with learning disabilities: strategy utilization and metacognitive awareness. It has been shown that children with learning disabilities employ inefficient and ineffective strategies (Hallahan & Bryan, 1981; Hallahan & Reeve, 1980; Torgesen, 1980). For example, Havertape and Kass (1978) recorded the vocalized self-directions of students with learning disabilities and average achieving students as they were attempting to solve problems. Relative to average functioning children, children with learning disabilities applied fewer appropriate strategies to problem solutions. Indeed, much of the time they appeared to implement strategies that were random and impulsive. In the

social domain Carlson (1987) found deficiencies in the ability of children with learning disabilities to generate appropriate strategies in conflict situations relative to those displayed by children without learning disabilities.

With respect to the issue of metacognition, other studies indicate that many children with learning disabilities have difficulty choosing and deploying effective strategies because of deficits in metacognitive (executive) processing. For example, it has been shown that they cannot monitor their comprehension of new information and experience difficulties in regulating their learning by planning, checking, revising, and evaluating their problem solving activities (Bos & Filips, 1982; Pressley & Levin, 1987; Wong, Wong, Perry, & Sawatsky, 1986).

In the present research, strategy training was provided by teaching students a procedure for generating appropriate strategies to problematic social situations. Metacognitive awareness was encouraged by having students monitor and evaluate their solution strategies in a group problem solving format (more detail regarding these procedures is provided in the Method section below).

In the present research three major hypotheses were tested. First, if the experimental program was effective in improving social relationships, then it should lead to a climate of greater social acceptance in the classrooms in which the program was implemented. Thus participation in the social skills program that involved coaching, role playing, and information sharing would lead to improved sociometric status for the class as a whole where the program was implemented. Second, if the strategy training aspect of the program was effective, then students in the experimental group should generate more appropriate responses to novel hypothetical social problem solving situations than the students in the untreated group. This hypothesis was based on the assumption that repeated practice in generating strategies in response to problematic social situations would enable students to develop more effective problem solving strategies that would generalize to novel role-play situations. Finally, if strategy training was the aspect of the program that led to improvements in social acceptance in the classrooms where the program was implemented, then it seemed reasonable to suppose that change in appropriate strategy use would predict change in sociometric status. Accordingly, the third hypothesis was that treatment-induced change in strategy use in the novel role-play situations would predict treatment-induced change in social acceptance.

Method

Subjects

The subjects in this study were students who attended a special school for severely learning disabled students. The criteria for entrance into the school are that students show evidence of a significant discrepancy between intellectual potential and achievement and that they have had a previous unsuccessful experience in a special class placement for students with learning disabilities. School regulations prevented us from obtaining standardized IQ test scores; however, we did have access to the intellectual assessment reports. Analysis of these reports indicated that all the students in the study were in the average range of intellectual ability. In addition, all children in the study were at least two years below expected age level in reading as indicated on the annual individual reading assessments administered by each student's teacher. In

these assessments, informal reading inventories such as the Jerry Johns Basic Reading Inventory (Johns, 1985) and the Burns and Roe Informal Reading Inventory (Burns & Roe, 1985) were used.

Each class in the school has approximately eight students. Teachers of similar aged students tend to work collaboratively and combine their students throughout the day. For this reason, two classes of eight students participated in the experimental group and two classes of eight students served as controls. Because of the small size of the school it was not possible to find two groups of students of exactly the same age. For this reason there was a statistically significant difference ($t(26)=7.87$, $p<.01$) in the average age of students in the experimental group (10 males and 2 females; 12.2 years, $SD=.32$) as compared with students in the control group (9 males and 6 females; 13.3 years, $SD=.42$). Sample size differences between the two groups (experimental group $n=12$; control group $n=15$) were because parents of four children in the experimental group and one child in the control group did not return the consent form. The consent form described the general purposes of the study and the evaluation instruments to be used. Parents of children in the experimental group were informed that all students would participate in the social skills program and that consent for assessment only was being requested. Parents of children in the control group were informed of the purposes of the study and the fact that consent for assessment was being requested. Analysis of reading scores obtained from school-administered reading tests indicated the children in the experimental group were functioning at a 2.9 grade level whereas students in the control group were functioning at a 3.7 grade level. Although this difference was significant ($t(26)=2.64$, $p<.02$) a subsequent analysis using age as a covariate indicated that the difference in reading scores between groups was attributable to age differences; that is, when age was introduced as a covariate, the analysis was nonsignificant ($t(25)=1.21$, $p>.2$). Analysis of covariance (using age as the covariate) on the math scores indicated no significant difference between groups ($t(25)=.66$, $p>.5$).

Procedure

The treatment subjects received the social skills program twice a week for 45 minutes over the course of six months (January to June). The two classes that formed the experimental group were combined for the social skills lessons. The control group received no formal social skills program. Based on interviews with the control teachers, social difficulties were dealt with as they arose. The teachers indicated that they often dealt with social issues by presenting the issue to the class as a whole so that students could offer their views.

In this intervention study a collaborative consultation research model was used that allows individuals with diverse expertise to work together toward helping others (Idol, Paolucci-Whitcomb, & Nevin, 1986; West & Cannon, 1988; West & Idol, 1990). Consultation models are designed to prevent and remediate learning and behavior problems as well as coordinate intervention programs (West, Idol, & Cannon, 1989). According to this model, individuals who are working together should negotiate their responsibilities and expectations, define the nature and parameters of the problem, generate and select intervention recommendations, evaluate the intervention, and redesign the intervention if needed (West & Idol, 1990).

In this study a clinical psychologist worked with teachers in the classroom. The clinical psychologist and the teachers determined their roles and responsibilities with respect to the social skills intervention program, defined the nature of the students' social problems, and worked together to develop, modify, and evaluate the social skills program.

As noted previously, there were two main components of the program: coaching/role playing and information sharing. In the coaching/role playing aspect of the program, the students were presented with situations (see Table 1) that were jointly decided on by the clinical psychologist, teachers, and students. The procedure described by Oden and Asher (1977) was adapted to the whole class intervention. Each situation was described orally to the class and the students were encouraged to give their views on how they would react to the social problem. Once everyone's ideas had been aired, students broke up into small groups in order to develop role-plays that demonstrated the problem and how they would cope with it. The presentation of the role-plays was then followed by a debriefing discussion in which the students were encouraged to evaluate and monitor the strategies that had been demonstrated in the role-plays. In giving guidance to the students regarding their responses to the role-plays, three considerations influenced the teachers and the psychologist. First, the students should develop active responses to situations rather than passively going along with whatever other children were doing in the situation. Second, children were coached to stop and think before acting. Third, they were encouraged to generate strategies and arrive at compromise solutions that would take all viewpoints of participants into consideration.

The information sharing aspect of the program initially took the form of self-awareness activities (i.e., how I see myself; how others see me; what are my interests? what are my strengths and weaknesses?) that were eventually shared with other students. Subsequently, students were given the opportunity to share their perceptions of each other (e.g., students were given the assignment of saying something positive about other children in the class) and to compare their self-perceptions with the perceptions that were reported by other students. The self-awareness/information sharing aspect of the program was derived from the Metacognitive Approach to Social Skills Training (Sheinker & Sheinker, 1988). In implementing this aspect of the program it became evident that students experienced difficulties in exchanging information because of a style of communication that was often excessively confrontive. To address this issue the students were divided into three groups that were each led by an

Table 1
Situations Targeted for Intervention

1. Knowing how to respond when the teacher asks a student why he/she is having difficulty reading.
2. Knowing how to ask for directions when you get lost.
3. How to approach a classmate who you would like to make friends with.
4. How to ask a person about his/her beliefs or culture when they are different from your own.
5. How to join a new club at a new school you are attending.
6. How to cope with a new school setting.

adult (the clinical psychologist and two classroom teachers). They were provided with specific exercises and were directed toward discussing issues (such as the difference between a fact an opinion) that would facilitate communication and information sharing.

Evaluation Measures

Checklist data. It is well established that 30-40% of children with learning disabilities have hyperactivity (Lambert & Sandoval, 1980). This fact is particularly relevant to the present study because of evidence that measures of hyperactivity are good predictors of peer sociometric ratings (Bruck & Hebert, 1982). Thus, prior to the start of the study, teachers were asked to complete one copy of the Abbreviated Conners Questionnaire (ACQ; Conners, 1969) for each student. The ACQ is usually used as an indicator of hyperactive behavior, and the 10 items on the scale consist of characteristic behaviors of hyperactive children (e.g., fidgety behavior, short attention span, temper tantrums). Each item is rated on a scale of 0 to 3: not at all, just a little, pretty much, and very much. The scores on the Conners items were summed for each student and this score was used to provide an index of hyperactive behavior.

Teachers also completed a copy of the Taxonomy of Problem Situations (TOPS; Dodge, McClaskey, & Feldman, 1985) for each of their students. The TOPS consists of 44 items that describe common problematic social situations (e.g., name calling, difficulties in turn taking) in a school context. Each item is rated on a five-point scale with 1 signifying that the behavior is never a problem and 5 that it is almost always a problem. In filling out the TOPS, teachers were instructed to score an item as a problem in terms of the appropriateness or inappropriateness of the child's response to the problematic situation. For example, a child would be rated as having a problem with name calling only if the teacher believed his or her reaction to name calling was inappropriate. For each student the responses to the 44 TOPS items were summed in order to provide an overall index of each student's social difficulties.

Students were asked to complete a self-perception questionnaire in a group-administered format that was identical to the TOPS except that the questions were reworded (by changing the pronouns from the third person to the second person) to make them answerable by the students themselves. In this case, students were instructed to indicate how problematic it was for them to react to the situation. As with the original TOPS questionnaire, each item is rated on a five-point scale with 1 signifying that the behavior is never a problem and 5 that it is almost always a problem. The students' responses were averaged and a total score was calculated as an index of each student's perceived social difficulties. In addition, the items from the adapted TOPS for students were used to construct the hypothetical situations as described below.

Hypothetical situations. In order to measure their strategy use, students were asked to respond to a number of hypothetical problem solving situations during individually conducted interviews. The hypothetical situations were derived from the items that appeared on the student version of the TOPS. In developing the hypothetical situations, the six student TOPS items that posed the most difficulty across all students in the study were used. The six items included: name calling; having a friend refuse an offer to play, and then seeing

the same friend play with someone else a short time later; having a group of classmates laugh at you when you do poorly at a game or activity; having friends leave you out of an activity; asking a friend for an object that you want; having to ask for something back that a friend has borrowed from you. These items from the student version of the TOPS were reworded into hypothetical problem solving situations. For instance, the hypothetical situation that addressed the problem of name calling was worded as follows: "If you attempted to join a group of other students and one of them said 'What do you want egg-head,' how would you respond?" Another scenario asked how a student would react if he or she had lent a possession to another student and the student was late in giving it back. In this case the student was asked to generate the name of a valued object that might be lent to another student. He or she was then asked, "If you lent (insert name of object) to a friend and he/she was late in giving it back, what would you do?" Responses to these situations were coded by trained raters with respect to seven categories derived from Carlson (1987) described in Table 2.

The first three categories (assertive, accommodative, and evaluative) were considered positive responses in light of the program goals to be actively involved, to stop and think, and to arrive at compromise solutions. The last three (passive, egocentric, and antisocial) were classified as negative. Students were presented with the hypothetical situations in two different ways. First, they were presented with the situation and asked what they would do. After responding to this question, they were given the same situation with a specific goal in mind (e.g., in the name calling situation they were asked to indicate what they would do to get the other child to stop calling them a name).

Table 2
Hypothetical Problem Solving Scoring Categories

Positive Strategies

1. *Assertive*
   Responses that deal with the problem in a direct, forthright manner. They allow the child to get what he/she wants while fostering or maintaining good social relationships
2. *Accommodative*
   Responses that are less direct and self-assured. They may not result in the child getting what he/she wants but will still foster or maintain good social relationships.
3. *Evaluative*
   Responses in which the child reflects on how he/she would feel or think in the situation. These include responses in which the child tries to see the situation from the other child's perspective.

Negative Strategies

4. *Passive*
   Responses that do not deal with the problem or seek social contact with others. The child will ignore, leave, or "not worry about it."
5. *Egocentric*
   Responses in which the child seeks to get his/her way at the expense of social relationships. This includes lying, stealing, cheating, nagging, whining, threatening, taking over the game, being bossy or belligerent.
6. *Antisocial*
   Retaliatory responses. The child does not seek the goal and only wants to repay other(s) for some perceived injustice. The child wants to hurt others either emotionally or physically.

Raters (the second author and a bachelor's level research assistant) who scored the students' social problem solving responses required approximately 10 hours of training in order to achieve an interrater reliability of .85 or greater. Reliability was calculated by dividing agreements by agreements plus disagreements of ratings obtained from two independent raters on 10% of the interviews that were conducted. One of these reliabilities was calculated before formal scoring was initiated. The second was conducted after half the interviews had been scored. The reliability scores were 88.2% and 92.1% respectively.

Sociometric evaluation. Sociometric ratings were obtained by having students rate all other children in the two classes that received the experimental program. The two classes were separate for most of the day, but came together for approximately six hours per week (for social skills instruction and for some language arts periods). Thus approximately half of the students who were rated by each student were familiar because they were in the same class for most of the day. The other half of the students were less familiar because they were in the same class for only six hours per week. The ratings were conducted in the same manner in the control classes. In this case children from the two control classes came together for approximately six hours per week for health and language arts. It is also important to note that there were only about 80 students in the school in which the study took place, so most students, especially those of similar age, were quite familiar with one another. Sociometric ratings were obtained in individual interviews in which subjects were given a deck of cards with the name of one student on each card. They were asked to sort the cards in response to three questions: (a) how much would you like to sit next to this person on the school bus? (b) how much would you like to hang out with this person on the playground? and (c) how much would you like to work with this person on a class project? The student was asked to sort the name of each child into one of three piles that were labeled: a lot, a little, not at all. For each child the ratings given by all other children of him or her were assigned a numerical code (a lot=1; a little=2; and not at all=3), summed and then entered into a data base for statistical analysis. Thus popular children would be those with relatively low scores; unpopular children would be those with relatively high scores.

Results

Checklist Data

Analysis of the TOPS completed by teachers indicated eight situations where the average ratings of teachers exceeded 3.0 (out of a possible 5). Student self-ratings on the adapted TOPS for students were generally lower with only two situations exceeding 3.0. A preliminary analysis was undertaken to determine if there were pretest differences between the experimental and control groups, on the teacher and student TOPS. This analysis revealed that the average teacher TOPS ($t(26)=.64$. $p>.1$) and student TOPS ($t(26)=.25$, $p.>1$) did not differ between groups.

Analysis of the pretest Abbreviated Conners Questionnaire responses indicated no differences between groups ($t(26)=.28$, $p>.5$; experimental group $M=8.75$, $SD=7.12$; control group $M=8.13$, $SD=4.09$).

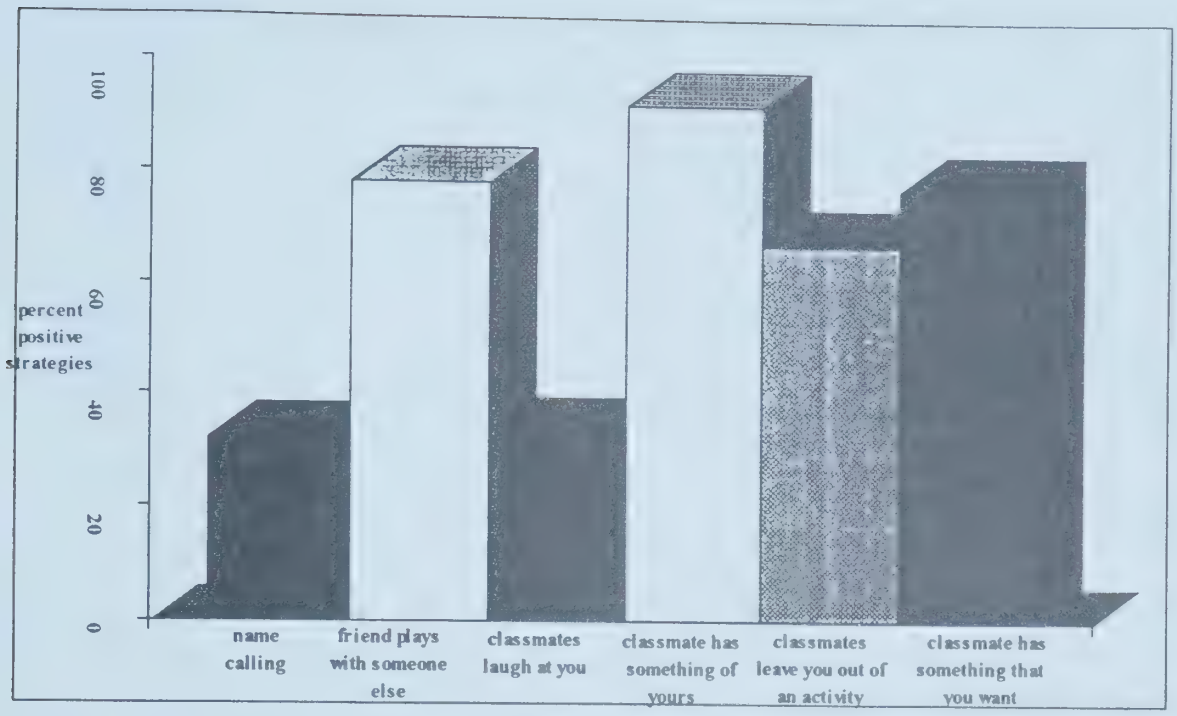


Figure 1. Percent positive strategies at pretest as a function of situational context.

Hypothetical Solution Strategies

As noted previously, students' categorized responses were reclassified into two solutions: assertive, accommodative, and evaluative categories were defined as positive solutions and passive, egocentric, and antisocial categories were defined as negative solutions. Pretest responses averaged across both experimental and control groups are shown in Figure 1, which indicates that there was considerable variation in the level of strategy use across situations. Two situations (name calling and being laughed at by classmates) showed a markedly lower incidence of positive strategies at pretest as compared with the other situations (χ^2 (5)=49.11, $p.<01$).

As noted previously, each hypothetical situation was administered in two ways. In the standard presentation condition, students were presented with the situation and then asked what they would do. In the goal-given condition, after the situation was presented students were given a social goal and then asked what they would do to achieve that goal. Between-group percent positive responses at both pre- and posttest to the hypothetical situations are shown in Table 3 for the standard presentation condition and in Table 4 for the goal-given condition. First, for the standard presentation condition, in order to determine the presence of pretreatment differences, between-group responses were compared using the Fisher Exact Probability Test. This analysis revealed that the responses of the experimental and control groups did not differ on any of the six situations. For the goal-given condition, there were pretreatment differences in favor of the experimental group for the situation in which "a peer has a toy, game, or object that I would like" ($p.<03$). Analysis of treatment effects were made in two different ways. First, posttest responses were compared using the Fisher Exact Probability Test. This analysis revealed that there were significant differences in only two situations: name calling, $p.<02$; and being laughed at by classmates in the goal given condition, $p.<03$. Second,

Table 3
Percent Positive Strategies in Responding to the Hypothetical Situations

	Experimental		Control	
	Pre	Post	Pre	Post
Name calling	25.0	60.7	40.0	21.4
Children laugh at you	36.4	54.5	33.3	40.0
You ask a child to play and find that they have played with someone else	73.3	60.0	83.3	66.7
Friends leave you out of an activity	58.3	58.3	72.7	90.9
A classmate has an object that you want	71.4	92.8	83.3	75.0
A friend has something of yours that you want back	93.3	86.7	90.9	90.9
Mean	59.6	68.8	67.2	64.1

analyses of the within-group change scores indicated significant positive change as assessed by a binomial test in the experimental group but not the control group for both of these situations: name calling ($p.<05$), and being laughed at by friends in the goal-given condition ($p<.05$). There were no significant change scores for any of the other situations in the experimental group and no significant effects for the control group for any of the hypothetical situations.

Sociometric Status

Change in sociometric status was assessed by assigning a score to each student based on the ratings received from the other students in the class (1=a lot; 2=a little; 3=not at all). Thus in this system a high score was indicative of a low status rating. The sociometric data are shown in Figure 2.

To evaluate sociometric scores a repeated measures split plot analysis of covariance with age as the covariate was performed with one between-group variable (GROUP: experimental vs. control) and two within-subject variables (SITUATION: with 3 levels: bus, playground, working on a class project and

Table 4
Percent Positive Strategies in Responding to the Hypothetical Situations in the Goal-given Condition

	Experimental		Control	
	Pre	Post	Pre	Post
Name calling	50.0	58.3	18.2	9.1
Children laugh at you	20.0	70.0	40.0	26.7
You ask a child to play and find that they have played with someone else	83.3	91.7	86.7	86.7
Friends leave you out of an activity	72.7	72.7	100.0	75.0
A classmate has an object that you want	100.0	90.0	58.3	83.3
A friend has something of yours that you want back	90.9	54.5	64.3	64.3
Mean	69.48	72.86	61.25	57.50

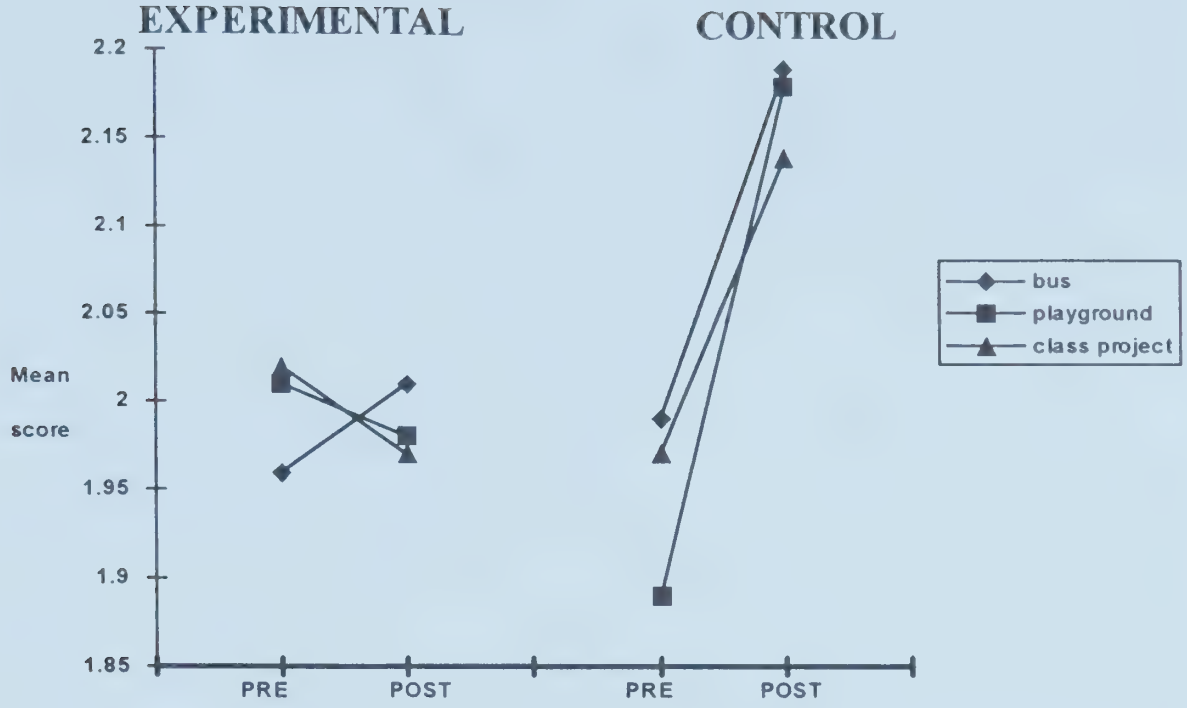


Figure 2. Social acceptance as a function of treatment conditions.

TEST SESSION: with 2 levels: pretest and posttest). The results of this analysis indicated that there was a significant main effect of group ($F(1,25)=13.60$, $p.<01$) indicating lower sociometric status in the control group and a group x test session interaction ($F(1,25)=7.47$, $p.<02$). Decomposition of this effect indicated that there was no change in the experimental group from pre- to posttest ($F(1,11)=.11$, $p>.5$), *but there was a significant decline in status in the control group* ($F(1,14)=19.63$, $p<.001$).

It was hypothesized that the beneficial effects of the program on social acceptance would be mediated by improvements in social problem solving strategies. If this hypothesis were true, then one would predict significant correlations between social problem solving and the sociometric measures at posttest when controlling for the pretest correlation between these two variables. To examine this issue, a partial correlational analysis was conducted on the social problem solving data for the two situations for which there were significant treatment effects (name calling and classmates laughing) and the sociometric "playground" measure. The results of this analysis indicated that from pre- to posttest there was greater stability for the sociometric measure ($r(25)=.87$) than for the problem solving measure ($R's$ (25)=.25 and .29 for the name calling and peer laughing situations respectively). Next, for both groups combined there were no significant correlations between responses to the hypothetical situations and sociometric status at pretest for either situation ($R's$ (25)=.08 and .01 for the name calling and peers laughing situations respectively). At posttest, these correlations increased to .27 and .24. The partial correlations were .30 and .21, neither of which is significant, suggesting that from pre- to posttest there was not a significant increase in variance in the sociometric measure accounted for by the social problem solving measure. Analysis of the separate groups indicated that in the name calling situation both experimental and control subjects showed an increase in the magnitude of the correlation

from pre- to posttest, but in neither case was that increase significant (partial $r(25)=-.35$ for the experimental group and .30 for the control group). For the peers laughing situation in the goal-given condition the correlation was stronger for the control group at pre- and posttest, but only the experimental group showed a significant increase in the value of the correlation from pre- to posttest (as indicated by the partial $r=.42$, $p<.05$). Thus for the one situation (peers laughing) it appeared that there was a program-induced increase in the relationship between problem solving strategies and sociometric measures that suggests generalization of the treatment effects.

Discussion

The results of the present study indicate that a social skills intervention program with emphasis on coaching, role playing, and information sharing conducted in an intact classroom can have some positive impact on the social acceptance and social problem solving skills of students with learning disabilities. In contrast, previous studies that have targeted students with learning disabilities for intervention and that have led to successful outcomes (see McIntosh et al., 1991) have been conducted in groups pulled out of their classrooms for intervention.

In terms of specific hypotheses, it was proposed that the intervention would have positive and significant effects on social acceptance. This hypothesis was not confirmed insofar as there was no positive change in the sociometric ratings in the experimental group. However, there was some evidence that the program was effective because there was a significant decline in sociometric ratings from pre- to posttest in the control group, but not the experimental group. This result may be a manifestation of a previous finding by Bierman et al. (1987) of a decline in the number of positive peer interactions over time in rejected boys who do not receive treatment. It may be that because there was no formal social skills program in the control group, the overall level of peer acceptance may have declined over the course of the year. A social skills program of the type employed in the present study may prevent such a deterioration in peer relationships. To verify this possibility, future studies should examine the relationship between the number of positive peer interactions and sociometric status in children with learning disabilities who are and who are not receiving social skills training.

Analysis of the data from the hypothetical situations indicated two situations that were especially problematic for the students in both groups: name calling and being laughed at by other students when performing poorly at a game. This result is consistent with evidence presented by Dodge, Petit, McClaskey, and Brown (1986) and Lochman and Lampron (1986) that situations involving provocation pose the greatest difficulty for aggressive children. The present study extends this finding for children with learning disabilities. In the present study these two situations were the only ones in which there were significant treatment effects: Children in the experimental group evidenced a significant improvement in the quality of problem solving strategies used at post- as compared to pretest in both of these situations. This effect suggests that the intervention was effective in improving the social problem solving skills of children in the experimental group, whereas no such effect was obtained in the control group. Carlson (1987) found that male children with learning dis-

abilities were more likely to select negative strategies than positive strategies to resolve conflict. Insofar as all the hypothetical situations used in the present study were conflictual in nature, the present study extends this finding by showing that intervention can increase use of positive strategies in conflict situations with children who have learning disabilities. It might be argued that the treatment effects on the hypothetical situations were weak because significant results were obtained from only two out of six situations. However, reference to Figure 1 shows that prior to treatment students used significantly lower-level strategies in the two situations where there were significant treatment effects. It seems reasonable to suppose that treatment effects may not have been realized in the other four situations because of ceiling effects.

It was hypothesized that the change in sociometric status would be mediated by improvement in social problem solving strategies. In the present study one significant relationship emerged; that is, when the pretest correlation between the sociometric measure and the social problem solving measure was removed by means of a partial correlation, there was a significant relationship between sociometric status and social problem solving for one situation only (peers laughing at the student). Thus there is some evidence in the present study that improvements in peer relationships were mediated by improvements in social problem solving skill. Because the relationship was quite weak and restricted to one situation, it is important to replicate this finding with another sample. The need for replication is underscored in this situation because in previous studies it has been the rule that social problem solving is not significantly related to measures of adjustment derived from teacher and/or peer perceptions (Dodge et al., 1986; Rickel, Eshelman, & Loigman, 1983; Sharp, 1981; Vitaro & Pelletier, 1991; Weissberg et al., 1981). The one exception is the study conducted by Gottman, Gonso, and Rasmussen (1975) in which a positive relationship was obtained between the sociometric measure and measures of both referential communication and knowledge of how to make friends.

Although the present study provides limited support for the hypothesis that a coaching/role playing/information sharing intervention can have impact on peer acceptance, the mechanism that mediates this effect is not clear. Although there was some evidence that the effects of the program on peer acceptance may have been mediated by changes in social problem skills, this effect was quite weak. The small magnitude of this effect may have been because cognitive and metacognitive development in students is multifaceted, requiring substantial time and effort in students and teachers. The only explicitly cognitive component of the intervention occurred during the debriefing sessions following the role-plays where students were encouraged to generate strategies, stop and think before they act, and monitor and evaluate their strategies. It may be that the connection between cognitive and metacognitive development and peer acceptance could be strengthened in future studies by targeting a greater variety of dimensions of cognitive and metacognitive strategies and by providing more depth in teacher training and intervention. It is also likely that in order to accurately measure the impact of a social skills intervention, it may be necessary to evaluate skill acquisition in real rather than hypothetical situations. This proposal is consistent with Vitaro and Pelletier's

(1991) finding that indicated weak relationships between performance in hypothetical and real social situations.

It is also possible that the magnitude of the treatment effects was limited by the fact that there may have been insufficient attention paid to student attributions. In the present study, the focus of attention was on strategy use and metacognitive awareness, but less so on attribution. Prior studies have suggested that many children with learning disabilities have motivational problems that stem from failure experiences and a tendency to attribute these failures to insufficient ability (Butkowski & Willows, 1980; Licht, 1983). Treatment programs that teach students to attribute past failures to insufficient effort tend to foster effort attributions and increased task persistence (Andrews & Debus, 1978; Dweck, 1975). Others have shown that effort feedback related to successful achievement enhances an individual's motivation and skills (Schunk, 1982). Borkowski and his colleagues have investigated the relationship between cognitive and motivational processes by examining the effects of attributional retraining and strategy training on children with learning disabilities (Borkowski, Weighing, & Carr, 1988). Their results showed that strategy training alone, without correcting negative attributional beliefs, is an ineffective method of instructing students with learning disabilities and is not likely to promote treatment generalization (Borkowski, Estrada, Milstead, & Hale, 1989). This finding underscores the importance of integrating metacognitive instruction and attributional retraining in programs designed to improve the generalized problem solving ability of children with learning disabilities. Future research should examine the impact of the executive functioning and attributional beliefs on the strategic behavior, generalized problem solving skills, and sociometric status of children with learning disabilities.

Training students with learning disabilities to become more effective and efficient social problem solvers is a multifaceted and challenging task that requires attention to cognitive, affective, and behavioral dimensions. Students need to be taught specific strategies and when and where to use them. Teachers need to facilitate the ability of students to monitor their own performance, develop skills in modifying and evaluating their strategic behavior, and provide them with feedback and practice. Moreover, intervention programs should be integrated in the school curriculum and include skills for children with learning disabilities that when applied will result in beneficial social outcomes such as enhanced feelings of self-worth and improved social status.

To date little research has been directed at investigating the relationship between the perceptions of social situations and strategy use in children with learning disabilities. Hence more research is needed to specify those social situations that children find difficult, why they find these situations difficult, and how this affects their social problem solving skills.

The results of this study provide some support for the utility of social interventions to ameliorate the social difficulties of children with learning disabilities, and extend the results of previous pull-out interventions to the whole classroom. However, comprehensive conceptual models that delineate the mechanisms by which social interventions affect the patterns of social acceptance of children with learning disabilities need further elaboration (Ladd & Mize, 1983; Pearl, 1987). Future research should examine the differential

effects of program components such as coaching, role playing, and information sharing on the social development of children with learning disabilities. In the present study, although some significant effects were obtained with the treatment, we are unable to discuss the relative contribution of the components of the program.

A number of methodological shortcomings to this study limit the conclusions to be made. First, because intact classrooms were used, it was not possible to randomly assign students to treatment conditions. Thus it is possible that the results were due to sampling bias. The only evidence against sampling bias is that on most measures that were taken (i.e., sociometric measures, social problem solving, academic functioning) there were no pretest differences. A second limitation of this study is that the control group was approximately one year older than the experimental group. To control for any effects of age, analyses of covariance were used whenever possible.

Although the results of this study must be considered tentative due to the above shortcomings, the study does expand on prior research on social intervention with children who have learning disabilities. One of the conclusions from this study is that the situation-specific social problem solving difficulties of children with learning disabilities should be more comprehensively delineated to better understand their social behaviors. This might allow for the design of treatments that can be individually tailored to meet the unique social needs of each child. This is consistent with the profile approach for identifying social skills and deficits proposed by Dodge and Murphy (1984) and the intervention approach of fitting social skills programs to target groups proposed by Coie (1985). Moreover, the present findings point to the need for more closely evaluating the relationship between perception of social difficulty and strategy use.

In the present program of research, it was evident from informal observations that limitations in children's expressive language skills influenced their ability to convey their ideas in coaching and role playing situations and to share information with their teachers and peers. Although expressive and receptive language skills were not included as variables in the research design, there is a need for future studies to evaluate social interventions with respect to children's communication skills across a range of social situations. This is consistent with Bryan's (1991) recommendation for research to link communicative and social competence to sociometric status.

This investigation did not evaluate whether the components of the social intervention program (coaching, role playing, and information sharing) were utilized by the teachers beyond the two days a week over the six-month period. Future studies should explore the generalization of the social learning process to the academic learning process.

References

Andrews, A., & Debus, R. (1978). Persistence and causal perception of failure: Modifying cognitive attributions. *Journal of Educational Psychology, 70*, 154-166.

Andrews, J. (1989). Social problems of learning disabled children: A review of the literature and some suggestions for assessment, intervention and future research. In C. Violato & A. Marini (Eds.), *Child development: Readings for teachers* (pp. 271-293). Calgary, AB: Detselig.

Association for Children with Learning Disabilities. (1985). ACLD description: Specific learning disabilities. *ACLD Newsbriefs*, 15-16.

Baum, D.D., Duffelmeyer, F., & Geelan, M. (1988). Resource teachers' perceptions of the prevalence of social dysfunction among students with learning disabilities. *Journal of Learning Disabilities, 21*, 380-381.

Bierman, K.L. (1986). Process of change during social skills training with pre-adolescents and its relation to treatment outcome. *Child Development, 57*, 230-240.

Bierman, K.L., & Furman, W. (1984). The effects of social skills training and peer involvement on the social adjustment of pre-adolescents. *Child Development, 55*, 151-162.

Bierman, K.L., Miller, C., & Stabb, S. (1987). Improving the social behavior and peer acceptance of rejected boys: Effects of social skill training with instructions and prohibitions. *Journal of Consulting and Clinical Psychology, 55*, 194-200.

Borkowski, J.G., Estrada, M.T., Milstead, M., & Hale, C.A. (1989). General problem solving skills: Relations between metacognition and strategic processing. *Learning Disability Quarterly, 12*, 57-70.

Borkowski, J.G., Weighing, R.S., & Carr, M. (1988). Effects of attributional retraining on strategy based reading comprehension in learning disabled students. *Journal of Educational Psychology, 80*, 46-53.

Bos, C., & Filips, D. (1982). Comprehension monitoring skills in learning disabled and average students. *Topics in Learning and Learning Disabilities, 2*, 79-85.

Bruck, M., & Hebert, M. (1982). Correlates of learning disabled students' peer interaction patterns. *Learning Disability Quarterly, 5*, 353-362.

Bryan, T. (1991). Social problems and learning disabilities. In B.Y.L. Wong (Ed.), *Learning about learning disabilities* (pp. 195-229). New York: Academic Press.

Burns, P.C., & Roe, B.D. (1985). *Burns-Roe informal reading inventory: Pre-primer to grade 12*. Boston, MA: Houghton Mifflin.

Butkowsky, I.S., & Willows, D.M. (1980). Cognitive-motivational characteristics of children varying in reading ability: Evidence of learned helplessness in poor readers. *Journal of Educational Psychology, 72*, 408-422.

Carlson, C.I. (1987). Social interaction goals and strategies of children with learning disabilities. *Journal of Learning Disabilities, 20*, 306-311.

Coie, J.D. (1985). Fitting social skills intervention to the target group. In B.H. Schneider, K.H. Ruben, & J.E. Ledingham (Eds.), *Children's peer relations: Issues in assessment and intervention* (pp. 141-156). New York: Springer-Verlag.

Conners, C.K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126*, 884-888.

Conte, R., & Andrews, J. (1993). Social skills in the context of learning disability definitions: A reply to Gresham and Elliott and future directions. *Journal of Learning Disabilities. 26*, 146-153

Dodge, K.A., & Murphy, R.R. (1984). The assessment of social competence in adolescents. In P. Karoly & J.J. Steffan (Eds.), *Adolescent behavior disorders: Current perspectives* (pp. 61-96). Lexington, MA: Heath.

Dodge, K.A., McClaskey, C.L. &, Feldman, E. (1985). Situational approach to the assessment of social competence in children. *Journal of Consulting and Clinical Psychology, 53*, 344-353.

Dodge, K.A., Petit, G.S., McClaskey, C.L. &, Brown, M.M. (1986). Social competence in children. *Monographs of the Society for Research in Child Development, 51*, 1-85.

Dudley-Marling, C.C., & Edmiaston, R. (1985). Social status of learning disabled children and adolescents. A review. *Learning Disability Quarterly, 8*, 189-204.

Dweck, C.S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *Journal of Personality and Social Psychology, 31*, 674-685.

Forness, S.R., & Kavale, K.A. (1991). Social skills deficits as primary learning disabilities: A note on problems with the ICLD diagnostic criteria. *Learning Disabilities Research and Practice, 6*, 44-49.

Fox, C. (1989). Peer acceptance of learning disabled children in the regular classroom. *Exceptional Children, 56*, 50-57.

Gorney-Krupshaw, B., Atwater, J., Powell, L., & Morris, E.K. (1981). *Improving social interactions between learning disabled adolescents and teachers: A child effects approach* (Research Report No. 45). Lawrence, KA: University of Kansas Institute for Research in Learning Disabilities.

Gottman, J., Gonso, J., & Rasmussen, B. (1975). Social interaction, social competence, and friendship in children. *Child Development, 46*, 709-718.

Gresham, F.M. (1988). Social competence and motivational characteristics of learning disabled students. In M.C. Wang, M.C. Reynolds, & H.J. Walberg (Eds.), *Handbook of special education:*

Research and practice, Vol. 2, Mildly handicapped conditions (pp. 283-300). New York: Pergamon Press.

Gresham, F.M., & Elliott, S.N. (1989). Social skills deficits as a primary learning disability. *Journal of Learning Disabilities, 22*, 120-124.

Gresham, F.M., & Nagle, R.J. (1980). Social skills training with children: Responsiveness to modeling and coaching as a function of peer orientation. *Journal of Consulting and Clinical Psychology, 18*, 718-729.

Gresham, F.M., & Reschley, D.J. (1986). Social skill deficits and low peer acceptance of mainstreamed learning disabled children. *Learning Disability Quarterly, 9*, 23-32.

Hallahan, D.P., & Bryan, T.H. (1981). Learning disabilities. In J.M. Kaufman and D.P. Hallahan, (Eds.), *Handbook of special education* (pp. 122-153). Englewood Cliffs, NJ: Prentice-Hall.

Hallahan, D.P., & Reeve, R.E. (1980). Selective attention and distractibility. In B.K. Keogh (Ed.), *Advances in special education, Vol. 1: Basic constructs and theoretical orientations* (pp. 141-181). Greenwich, CT: JAI Press

Havertape, J.F., & Kass, C.F. (1978). Examination of problem solving in learning disabled adolescents through verbalized self-instructions. *Learning Disability Quarterly, 1*, 94-99.

Idol, L., Paolucci-Whitcomb, P., & Nevin, A. (1986). *Collaborative consultation*. Austin, TX: Pro Ed.

Interagency Committee on Learning Disabilities. (1987). *Learning disabilities: A report to the U.S. Congress*. Washington DC: U.S. Department of Health and Human Services.

Johns, J. (1985). *Basic reading inventory: Pre-primer to grade eight*. Dubuque, IA: Kendal Hunt.

Ladd, G.W. (1981). Effectiveness of a social learning method for enhancing social interactions and peer acceptance. *Child Development, 52*, 171-178.

Ladd, G.W., & Mize, J. (1983). A cognitive-social learning model of social skills training. *Psychological Review, 90*, 127-157.

Lambert, N., & Sandoval, J. (1980). The prevalence of learning disabilities in a sample of children considered hyperactive. *Journal of Abnormal Child Psychology, 8*, 33-50.

Licht, B.G. (1983). Cognitive-motivational factors that contribute to the achievement of learning disabled children. *Journal of Learning Disabilities, 16*, 483-490.

Lochman, J.E., & Lampron, L.B. (1986). Situational social problem solving skills and self-esteem of aggressive and non-aggressive boys. *Journal of Abnormal Child Psychology, 14*, 605-617.

McIntosh, R., Vaughn, S., & Zaragoze, N. (1991). A review of social interventions for students with learning disabilities. *Journal of Learning Disabilities, 8*, 451-458.

Oden, S., & Asher, F.R. (1977). Coaching children in social skills for friendship making. *Child Development, 48*, 495-506.

Pearl, R. (1987). Social cognitive factors in learning disabled children's social problems. In S.J. Ceci (Ed.), *Handbook of cognitive, social, and neurological aspects of learning disabilities* (vol. 2, pp. 273-299). Hillsdale, NJ: Erlbaum.

Pressley, M., & Levin, J.R. (1987). Elaborative learning strategies for the inefficient learner. In S.J. Ceci, (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (vol. 1, pp. 175-212). Hillsdale, NJ: Erlbaum.

Rickel, A.U., Eshelman, A.K., & Loigman, G.A. (1983). Social problem solving training: A follow-up study of cognitive and behavioral effects. *Journal of Abnormal Child Psychology, 11*, 15-28.

Schneider, B.H. (1992). Didactic methods for enhancing children's peer relations: A quantitative review. *Clinical Psychology Review, 12*, 363-382.

Schumaker, J.B., & Ellis, E.S. (1982). Social skills training of LD adolescents: A generalization study. *Learning Disability Quarterly, 5*, 409-414.

Schumaker, J.B., & Hazel, J.S. (1984). Social skills assessment and training for the learning disabled: Who's on first and what's on second. *Journal of Learning Disabilities, 17*, 422-430.

Schunk, D.H. (1982). Effects of attributional feedback on children's perceived self-efficacy and achievement. *Journal of Educational Psychology, 74*, 548-556.

Sharp, K.C. (1981). Impact of interpersonal problem-solving training on pre-schoolers' social competence. *Journal of Applied Developmental Psychology, 2*, 129-143.

Sheinker, J., & Sheinker, A. (1988). *A metacognitive approach to social skills training*. Rockville, MD: Aspen Publications.

Torgesen, J.K. (1980). Conceptual and educational implications of the use of efficient task strategies by learning disabled children. *Journal of Learning Disabilities, 13*, 19-26.

Vaughn, S. (1985). Why teach social skills to learning disabled students? *Journal of Learning Disabilities, 18*, 588-591.

Vaughn, S., & La Greca, A.M. (1988). Social skills of LD students: Characteristics, behaviors and guidelines for intervention. In K. Kavale (Ed.), *Handbook in learning disabilities* (pp. 123-140). San Diego, CA: College Hills.

Vitaro, F., & Pelletier, D. (1991). Assessment of children's social problem solving skills in hypothetical and actual conflict situations. *Journal of Abnormal Child Psychology, 19,* 505-518.

Weissberg, R.P., Gesten, E.L., Rapkin, B.D., Cowen, E.L., Davidson, E., Flores deApodaca, R., & McKim, B. (1981). Evaluation of a social-problem-solving training program for suburban and inner-city third-grade children. *Journal of Consulting and Clinical Psychology, 49,* 251-261.

West, J.F., & Cannon, G.S. (1988). Essential collaborative consultation competencies for regular and special educators. *Journal of Learning Disabilities, 21,* 56-63.

West, J.F., & Idol, L. (1990). Collaborative consultation in the education of mildly handicapped and at-risk students. *Remedial and Special Education, 11,* 22-31.

West, J.F., Idol, L., & Cannon, G. (1989). *Collaboration in the schools: An inservice and preservice curriculum for teachers, support staff, and administrators.* Austin, TX: Pro-Ed.

Whang, P. L., Fawcett, S.B., & Matthews, R. (1984). Teaching job-related skills to learning disabled adolescents. *Analysis and Intervention in Developmental Disabilities, 4,* 29-38.

Wiener, J. (1987). Peer status of learning disabled children and adolescents: A review of the literature. *Learning Disabilities Research, 2,* 62-79.

Wong, B.Y.L., Wong, R., Perry, N., & Sawatsky, D. (1986). The efficiency of a self-questioning summarization strategy for use by underachievers and learning disabled adolescents. *Learning Disability Focus, 2,* 20-35.

The Alberta Journal of Educational Research *Vol. XLI, No. 1, March 1995, 103-121*

John A. Ross
Ontario Institute for Studies in Education

and

J. Bradley Cousins
University of Ottawa

Giving and Receiving Explanations in Cooperative Learning Groups[1]

This qualitative study examined patterns of giving and receiving explanations in three samples of students working in cooperative learning groups (N=65 grade 7/8, 45 grade 9/10, and 51 grade 9/10) on problem solving tasks (correlational reasoning problems). Students were randomly assigned to mixed ability cross-sex groups of four or pairs; the treatment was Slavin's STAD. A gap was observed between student practice and an ideal model of explanation exchanges. Students infrequently sought explanations. They gave explanations that were inadequate because they were uncertain about performance, lacked a language for describing their thoughts, had deficient teaching skills, and were insensitive to the needs of help seekers. Explanation givers did not monitor recipients' understanding of explanations. The article proposes that the benefits of cooperative learning will be enhanced if teachers teach students how to ask for and give quality explanations.

Cette étude qualitative a examiné les tendances qui ressortaient lorsqu'on donne ou lorsqu'on reçoit des explications en se basant sur trois échantillons d'élèves qui travaillaient en groupe dans le contexte de l'apprentissage coopératif en groupes (N=65 $7^e/8^e$ année, N=45 $9^e/10^e$ année, et N=51 $9^e/10^e$ année) sur des tâches de résolutions de problèmes (raisonnement de problèmes corrélationnels). En pigeant les élèves au hasard, on les sépara en paires ou en groupes de quatre ayant des habiletés mixtes et des deux sexes à la mode de Slavin selon le traitement STAD (Student-Teams-Achievement-Divisions), c'est-à-dire, en regroupant les élèves en équipes qui visent à l'atteinte d'objectifs particuliers. On observa un écart entre la pratique actuelle des élèves et un modèle idéal des échanges des explications possibles. Les élèves recherchaient rarement des explications. Ils donnaient des explications qui étaient inadéquates parce qu'ils étaient incertains de leur performance, n'avaient le langage approprié pour décrire leurs pensées, étaient déficients en habiletés pédagogiques, et étaient insensibles aux besoins de ceux qui sollicitaient de l'aide. Les élèves qui donnaient des explications ne surveillaient pas le niveau de compréhension de ces explications de ceux qui les recevaient. Cet article propose que les bénifices de la démarche coopérative seraient accrus si les enseignant(e)s enseignaient aux élèves à demander des explications et à et donner des explications de qualité.

The constructivist revolution in instruction (Posner, Strike, Hewson, & Gertzog, 1982) has stimulated interest in teaching methods in which students learn from one another. This article reports a qualitative investigation that

---

John Ross is a professor of curriculum and Head of the OISE Trent Valley Centre. His research interests are curriculum development, program evaluation, and student assessment.
J. Bradley Cousins is an associate professor. His research interests are knowledge utilization, organizational learning, and evaluation.

examined the explanations that students give and receive when they are working together to learn a new skill.

Theoretical Framework

The Importance of Explanations

Process-product studies of student-student interactions (the most complete reviews are in Webb, 1989, 1992) have consistently found that students who give explanations learn more than those who do not, even after student ability is controlled. Learning occurs because giving explanations requires the reorganization of the material to be learned (Bargh & Schul, 1980); it makes unconscious thoughts explicit (King, 1989, 1990), and it contributes to conceptual conflict that can lead to cognitive restructuring (Doise & Mugny, 1979). In contrast, receiving an explanation has a variable impact, regardless of whether or not an explanation is received in response to a specific request. Webb (1989) observed:

> even receiving help at what appears to be a sufficiently high level of elaboration is not often sufficient for high achievement.... Student interaction would have to be examined more closely to determine whether the explanations were relevant, understood, and applied and internalized by the target student. (p. 28)

Webb's suggestion focuses attention on the adequacy of explanations that are given. The low impact of explanations might also be related to the failure of help givers to oversee the recipients' use of explanations. Hooper (1992) found that the achievement of those who were given help was higher when the help givers checked recipients' understanding, but other studies have found that students do not often monitor each other's performance (Azmitia, 1988; Ross, 1994).

Another reason for the unproductive outcomes of receiving an explanation might be that students do not ask for explanations when it is appropriate to do so. Students do not know when they do not understand (Markman, 1977, 1979) and do not seek help when it is needed (Good, Slavings, Harel, & Emerson, 1987; Karenbenick & Knapp, 1988; Newman, 1990). Requests for explanation are infrequent in most studies. The highest means were reported by Webb and Kenderski (1985). They found that a grade 8-9 subsample of high ability girls averaged nine requests for explanation per 45-minute period. The means reported by other studies, and for the remaining students in Webb and Kenderski (1985), were dramatically lower. For example, Hertz-Lazarowitz (1989) found that students in grades 3-8 engaged in high-level discussions less than 2% of the time; Deering and Meloth (1990) found that 91% of academic help interactions were at an information exchange level. Kempa and Ayob (1991) found that only 10% of information exchanges of able 16-year-olds were at the explainer level; and Corno (1989) observed that fewer than 5% of the interactions of 5th graders involved instructing.

It is also possible that students do not know how to ask for explanations. Several studies have reported that characteristics of the question (such as being repeated, specific, and directed toward a particular individual) determine whether it will be answered (Wilkinson & Calculator, 1982; Wilkinson & Spinelli, 1983).

These findings suggest that receiving an explanation is a productive form of cognitive engagement only if three conditions are met. First, explanation seekers must recognize a need for help, diagnose its specific dimensions, locate a suitable help giver, and access the peer resource through an appropriate request. Second, the explanation seeker must receive a quality explanation, consisting, for example, of the peer modeling the correct performance while providing an oral commentary that describes how the performance was generated, using language and ideas familiar to the learner. The specific teaching moves of effective explanation givers may vary, approximating techniques such as participant modeling instruction (Corno, 1992), cognitive apprenticeship (Brown, Collins, & Diguid, 1989), and expert scaffolding (Palincsar & Brown, 1984). A quality explanation is one that gives the specific answer required by the immediate task while providing the student with sufficient procedural knowledge to solve problems of a similar type on his or her own in the future. Third, the explanation seeker needs to internalize the explanation by applying it to the learning task and obtaining corrective feedback.

Cooperative Learning

Cooperative learning approaches differ in the attention they give to training students in group interaction skills. For example, Johnson and Johnson (1987) suggest that practice and reinforcement of peer tutoring skills are essential. Their approach provides an extensive array of activities to identify, demonstrate, and consolidate skills prerequisite to students learning from one another (Johnson, 1990). Feedback on cooperative skills is also a critical component of group investigation (Sharan & Shachar, 1988) and of the eclectic approach of Kagan (1988). There is evidence that prosocial behaviors increase in these approaches (Johnson, Johnson, & Stanne, 1986), although the unique contribution of the social skills component has not been disentangled from other treatment dimensions.

In contrast, the methods of the Slavin research group (Teams-Games-Tournaments, Student-Teams Achievement-Divisions [STAD], Team Accelerated Individualization, and Cooperative Integrated Reading and Composition) provide for no explicit instruction in group interaction skills, and formal student feedback focuses on cognitive achievement alone. Teachers are advised, in STAD for example, to praise teams that are working well, but specific peer interaction skills are not negotiated with students or otherwise highlighted (Slavin, 1990). Slavin's reviews of the literature persuade him that the combination of individual and group rewards through the fair allocation of extrinsic rewards is sufficient to promote learning (Slavin, 1983, 1987, 1990, 1992).

> Students may be motivated to engage in elaborated, cognitively involving explanations and discussions if the learning of their group-mates is made important by the provision of group rewards based on individual learning performances. (Slavin, 1987, p. 1166)

> The evidence supports a conclusion that group member helping on a group task and group member norms supporting performance are consequences of cooperative incentive structures. (Slavin, 1992, p. 153)

Research Questions

The specific purpose of the research reported here was to observe students working in STAD groups. We wanted to find out if the high-level helping predicted by Slavin would be visible. Because no previous study has examined explanation exchanges in STAD groups, we were initially guided by the three conditions for effective exchanges described above. We selected a qualitative approach because we believed that process-product research on explanation giving and receiving has been limited in four ways. First, the quantitative approach breaks students' speech into discrete utterances threatening the coherence of the whole (Solomon, 1992). Second, most of the research has been conducted in settings that gave little support for student helpfulness. For example, very few of these investigations used reward interdependence to encourage explanation givers to attend to the needs of students who needed help or instituted individual accountability to reduce social loafing. Third, the tasks assigned to students were rarely challenging: learning routine algorithms, for example, is unlikely to stimulate high-level cognitive exchanges. Fourth, little attention (with the exception of Hooper, 1992 and Webb, 1992) has been given to students' use of the explanations they receive.

Method

Sample

Productive conditions for studying explanation exchanges were provided in school districts that were responding to changes in Ontario curriculum policy that required that students acquire correlational reasoning skills in grades 7-10 (Ontario Ministry of Education, 1988). Previous research has found that correlations are the most challenging of the formal reasoning tasks (Lawson & Bealer, 1984; Yap & Yeany, 1988). Coincidentally the province was promoting the implementation of cooperative learning in the same grades as a mechanism for dealing with the elimination of ability tracking in the formation of classes (Ontario Ministry of Education, 1992).

There were three samples: (a) 65 grade 7 and 8 students working in groups of four in four classrooms using STAD; (b) 45 grade 9 and 10 students in three classes in a similar setting; and (c) 51 grade 9 and 10 students in three classes with each pair sharing a computer following a STAD approach. The average age was 13 years for sample (a) and 15 years for samples (b) and (c). Students were drawn from schools in four districts in central Ontario. Most schools were in small cities and towns. None was from a large city and there were few nonwhite students in the samples. Males and females were equally represented except in sample (b) which had slightly more females.

Instruments

Student interactions while in cooperative groups were audio- [samples (b) and (c)] or videotaped [sample (a)] on a single occasion at the end of the program for 25-45 minutes (the average was 30 minutes). The qualitative analysis of these data is reported in this article. In addition, quantitative data (student achievement and attitudes toward helpfulness) were collected. These data are not reported here; findings are described in Ross and Cousins (in press).

Treatment

In designing the treatments STAD procedures were followed closely. Materials were prepared using direct instruction modules that produced high levels of achievement in earlier experiments (Cousins & Ross, 1993; Ross & Cousins, 1993a, 1993b). A procedure for solving two variable and multivariate correlational problems was demonstrated by the teacher. Students were given practice activities that required the construction of graphs using content from the grade 7/8 or grade 9/10 geography curriculum. Students were assigned to teams as detailed in Slavin (1990): Teachers rank ordered students from least to most capable, using students' past performance on tests and observations of student behavior. The list was divided into ability quartiles and one student was randomly assigned to each group from each quartile. The resulting teams were adjusted by teachers to ensure a balance of genders and ethnicities and to avoid delinquent alliances (replacements were made within ability quartiles). Students in samples (a) and (b) completed practice activities in mixed-sex, mixed-ability groups of four. Students in sample (c) completed the same activities in mixed-sex, mixed-ability pairs; each pair shared a computer, using the software program CORReoGRAPH.[2]

In all treatments students worked in cooperative groups for 70% of the 300-minute duration of the program. During each practice period students were given two exercises to work on. The student task consisted of constructing scatterplots and trend lines to represent the relationship between two continuous variables, in some instances while controlling for a third variable. Figure 1 provides an example of a correlational reasoning task; Figure 2 provides an example (produced by CORReoGRAPH) of an answer to the exercise. Students in samples (a) and (b) completed all activities by hand; students in sample (c) did a few examples by hand, but most were created by students on the computer.

Each student was expected to complete his or her own graphs, asking for and receiving help when it was required. Near the end of each period teachers distributed answers to the exercises. Students were told that the purpose of the group activities was to learn how to solve correlational problems that would be assessed in two quizzes. Individuals completed each quiz without assistance from team members. Reward structures combined individual and group accountability using Slavin's (1990) fair sharing procedure that makes individual performance visible and combines each group member's improvement scores in calculating group rewards. The rewards varied between classes but all were extrinsic, most frequently consisting of team certificates, stickers and extra computer time. Students stayed in the same groups for the duration of the correlational reasoning module.

The treatments differed in one significant way from the original STAD. Although students had been working in groups previously, they had not been placed in structured STAD teams. Consequently the duration of the treatment was less than what is recommended for full STAD implementation.

Procedure

Verbatim transcripts of the audio/videotapes were made and read in their entirety by members of the research team. Attention focused on the explanation exchange protocols (which included interactions in which explanations

People Make Garbage

Angel and Tamil both live near a landfill site and they are very concerned about the amount of garbage that is produced in their community.

Angel: *Some people just don't care. The family next door to us puts out at least 8 bags of garbage every week. There ought to be a law that nobody can put out more than 3 garbage bags per week.*

Tamil: *Not so fast, Angel. What matters is the weight of garbage, not how many bags it is in. And what about the size of the family? Large families probably produce more garbage than small families.*

Angel: *I don't think that's true about the size of the family. What counts is whether they are recycling.*

Angel and Tamil decided to investigate the question. They collected information on the weight of garbage put out by each household. They recorded the number of people in the household and recycling habits (whether or not the household used a blue box and a compost pile).

Use the information collected to find out if there is a relationship between size of household and amount of garbage produced.

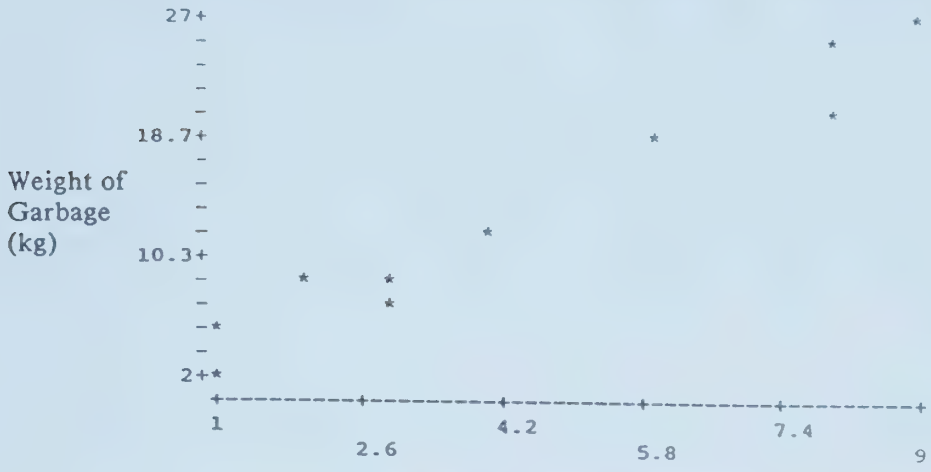
Household	*Number of Members*	*Weight of Garbage (Kg)*	*Recycling Habits*
Armitage	2	8	no
Balson	6	18	no
Carlson	3	2	yes
Duncan	9	27	no
English	2	3	yes
Fuller	1	2	no
Gillespie	4	2	yes
Horvath	7	6	yes
Ibey	10	8	yes
Johnstone	1	5	no
Kay	3	3	yes
Lawler	8	5	no
Munson	2	6	yes
Newman	3	1	yes
O'Toole	4	12	no
Porter	2	2	yes
Roberts	4	4	yes
Scott	3	7	no
Thompson	3	8	no
Urbach	8	21	no

Figure 1. Example of correlational reasoning task.

were offered without an overt request). The request-explanation-practice model described above provided the general framework, and specific issues evolved as we made multiple passes through the data. The final set focused on the clarity, precision, and courtesy of the request, the requester's persistence, and his or her selection of a potential help giver. In inspecting the explanations that were offered, we focused on the fit between the explanation and the specific need of the request, the degree of elaboration, the suitability of examples that were used, the extent to which the explanation giver diagnosed the needs of the help seeker, the number of students involved in giving the ex-

GRAPH A

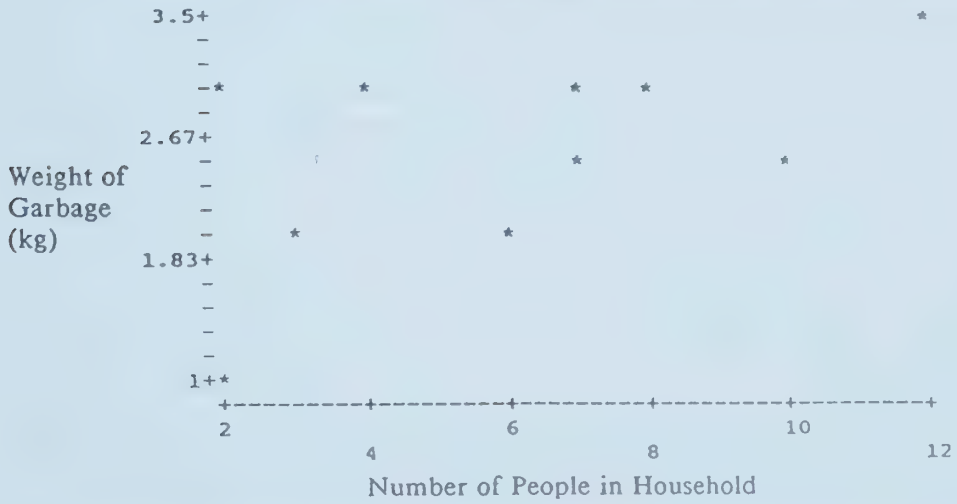
HOUSEHOLD SIZE AND GARBAGE PRODUCED FOR NON-RECYCLERS



For non-recycling households, there is a positive correlation. As the number of people in the household increases, the amount of garbage increases.

GRAPH B

HOUSEHOLD SIZE AND GARBAGE PRODUCED FOR RECYCLERS



For recycling households, there is a positive correlation. As the number of people in the household increases, the amount of garbage increases, but not to the same degree as in non-recycling households.

Figure 2. Example of answer to correlational reasoning exercise.

planation, the accuracy of the explanation, and the affective climate of the interaction. In examining instances of monitoring we focused on whether help givers asked if the explanation was understood, any tests of the help seekers' grasp of it, the assignment of practice activities, and attempts to reexplain. Although the samples were examined separately, the interpretive framework was shared and the results were highly similar. The findings are reported for the three samples combined, except where differences require a separate discussion.

Results

Explanation seeking. Only a third of the students requested an explanation and only half of these students received one in response. The requests for explanation were all sincere; none represented unnecessary help seeking. But requests tended to be poorly articulated. Few were directed toward a particular individual, most were brief, and many were vague, for example, "What are you doing?" In contrast, a student who saw that his graph did not have the same shape as one drawn by another group member, posed a much more precise request: "Since this [the distribution of data points] is much more spread out than yours, since I went up by ones [in scaling the axes], would [the teacher] still consider this a positive correlation?" Help seekers tended to be persistent, paraphrasing requests that were unanswered, and returning to the issue if they were unsatisfied. Requests that were specific and repeated were more likely to receive a response. Students who were unpleasant to their peers were less likely to receive assistance, and requests to students who were not members of the same pair or group were more likely to be ignored.

Explanation giving. There were several examples of explanations that were effective and brief. For example, a student who asked her partner why he was plotting different numbers than she was was told: "I'm doing the yes graph, that's why," meaning that he was plotting the same relationship (between family size and amount of household garbage), but for a different value of the control variable (recycle, yes or no). Similarly, a student who asked a peer why she was scaling the temperature variable in her graph from what appeared to be high numbers to low numbers, instead of the recommended procedure of scaling from low to high, was told "'cause negative 12 is colder than negative 1." There were also good longer explanations. In the following passage, Rachel [pseudonyms are used throughout] wanted to know why her graph did not look like that of another group member, in particular why she had more scale intervals on the bottom axis.

Rachel: What does yours look like, how many do you need across?

Loughery: I'll need a different number than you.

Rachel: Why?

Loughery: Because you're using different numbers than me. Your degrees only range from 4 to 9; mine go from 9 to 18. [i.e., Rachel's intervals are 5 degrees wide; Loughery's are 9 degrees].

Some students gave explanations that combined an oral commentary with a nonverbal demonstration. For example,

Sean: I don't know how to do these things.

Wesley: Okay. So we're doing the no's right? [i.e., they are plotting the households which do not recycle their waste]. So circle all your no's.... Now we go, okay, 2 and 8 [i.e., a household with 2 members producing 8 kg of garbage per week] right? So's all you have to do is go [up the y axis to] your 2, you go up 2, across [on the x axis to] 8. Right about there. ding, ding, ding

Students sometimes referred to a past example when giving a demonstration.

Ryan: How do you do it?

Stephanie: I'm not sure. [Pause] You guys. For this next one, I do the same as I did for this one? Okay guys?

Brandy: Sure.

Stephanie: So number of members is on the vertical axis and it's going up by 1, 2, 3, 4. And the horizontal axis is by 2s.

Some student explanations were probably too brief to be helpful to the help seekers. When Christa asked why all the data for the three variables given in the exercise could not go on the same graph, Natalie's explanation was that "it would be inconclusive if we did it like that." There were also explanations that contained misconceptions that probably confused the recipients. In the following Nate tried to describe how to distinguish between a negative and positive correlation.

Nate: That's positive. You know how you can tell? This way. As the graph goes up, as these numbers go up, these numbers go up too. If it was going like this, as these numbers go up, these ones would go. You know, how like, two positive make a

Valerie: negative.

Nate: yes.

Valerie: Positive. Two positive make a negative.

Nate: yes.

Valerie: Two negatives make a positive ...

Nate: And a positive plus a negative makes a negative ... as you're going up this way, both of these are going up, so it's positive [correct definition of a positive trend]. If you are going down that way, as this number goes up, this one would get lower, on here [correct definition of negative trend]. See? So a negative plus a positive times a positive will make a negative.

Valerie: Oh.

About 30% of the explanations contained errors. Some misconceptions in the explanations could have arisen from the models given in the lessons. For example, Christa believed that correlational predictions always had to be in the form recommended in the lesson, "if the temperature increases, then the snowfall will decrease," and that expressions like that of Wanda, "if the temperature decreases, then the snowfall will increase," were incorrect. Several students believed that when solving a problem involving more than two variables, the one that was in categorical form (e.g., inland/coastal location) should be selected as the control variable rather than making the decision on the basis of conceptual analysis.

Almost half the explanations appeared to be insensitive to the requestor's level of understanding. Some students persisted with explanations based on examples provided by the teacher even when it was inappropriate to do so. In the episode below Rachel did not understand why the group needed to produce four graphs when she thought three would do. The passage is particularly interesting because it shows how two students (Wanda and Loughery) worked together to help her even after she indicated that she should have known how to do it. Their intervention was unsuccessful; Rachel insisted later in the session that only three graphs were required.

Rachel: How do you get 4 [graphs]?

Wanda: You know how we did the soya bean one where you had to do, where there was degree days and averages.... Well, we have to figure out which one's the deciding factor for this one ... [when this example did not persuade Rachel, Wanda tried another]. You know when you did the cooks you grouped them all together and

Rachel: I wasn't here.

Loughery: Yes you were; [the teacher] showed it up on the overhead.

Wanda: Yesterday.

Rachel: Okay. I wasn't paying attention.

Wanda: Well maybe if you paid attention this wouldn't happen.

Loughery: Rachel, do you understand how he did that?

Rachel: No I don't.

Wanda: Remember when he put it on the overhead and he put all the cooks together and there was no correlation?

Rachel: Listen, just tell me what I have to do.

Wanda: You have to understand it to know what to do [she then rehearsed the teacher's lesson].

Explanation givers often confused help seekers, and at times themselves, by using different contexts to explain the procedures and by moving quickly from one instance to another. For example, Jacob was uncertain about how to set the scale for the axes in his graph. Jackie tried to help by demonstrating the skill, but she selected a different axis than the one he was working on and did not explain how she selected the upper end of the scale. She then tried to recall the example from the lesson without success.

Jacob: Now what am I going to go [up] by? 200?

Jackie: How [do you] break [it]? I forgot. I think you're supposed to put another break [i.e., scale interval] in it right like this.

Jacob: Just put another break after that? Why, what have you got there?

Jackie: Put the break there because it goes up to 20 [which was the value of the highest case in the data set].

Jacob: 20, 21, and then it goes out there.

Jackie: So where do you put another break?

Jacob: You can just keep going.

Jackie: No. You know how [the teacher] puts a break way up here?

Jacob: No I didn't.

Jackie: On the board he did.

Jacob: No he didn't. He just went up to a 1000 and put a 1000 there.

Jackie: Is that what I should do?

Jacob: I don't know. I don't even know what I'm doing.

Jackie: I'm pretty sure that's what you do.

Some opportunities for learning were also missed by help givers. For example, in one of the exercises students were given data on elevation, temperature, and latitude of weather stations on three continents. The stations were in the same latitude range on one continent and varied substantially in the other two data sets. They were asked to find the relationship between elevation and

temperature in each continent and to explain any differences they might find. In two of the groups students asked what they should do with the latitude data; in each case they were advised by their peers to ignore it.

Nathaniel: This is one of those things that they say. You're given more information than you need. We don't need this [latitude data]. Because [it's] just temperature and elevation.

On another occasion a student noticed that the pattern of data points in his graph was not the same as the pattern in the graph of his partner (because they used different scales). His partner was not interested in exploring why there might be differences.

Jason: What are you doing with them down here? Here we are up here.

Sandy: Well, that doesn't matter.

Many of the requests for explanations were met with off task behavior or the other members of the group continued with their own agendas. In at least one instance a request was ignored so that the group could finish: it was quicker to do it themselves than to explain it to someone else.

Wesley: Why don't you let me do it?

Gina: (to another group member) Come on, get to work. Hurry up so we can say we are done.

Requests for explanation were often met with procedural directives that dealt with the how, but not the why. For example, Ruth could not understand why the upper end of the scale in her partner's graph was so unlike her own. The partner responded to her request by describing his procedure, without explaining that he put a different variable on the axis.

Ruth: Why is it so small? That's not the highest [case].

Byron: We're doing this, the bottom one, by 3s; it works.

Explanation monitoring. Explanation givers were rarely observed testing the effects of their explanations. All instances of monitoring occurred in the grade 9/10 samples ([b] and [c]). One of the few examples consisted of a student (Anna) giving a lengthy recapitulation of the steps she and her partner (Randy) went through to complete one problem as a way of explaining how to do items of a similar type. The recitation appeared to be insufficient, but when Anna demanded that Dale attempt his own account of what they did he was able to do so succinctly.

Dale: [after the explanation] I still don't understand.

Anna: We did it right. We did it right.

Dale: OK, I think I understand now.

Anna: [skeptically] Explain it.

Dale: I can't. [but then he does] We did these two [pointing to household size and energy use, the variables to be correlated] with the no's [for the households in which there was no recycling]

Anna: Yeah.

Dale: And we did these two again with the yes [i.e., they correlated the same variables for the households in which there was recycling]

Anna: Yes

Dale: Oh, okay. I understand.

There was little co-construction of explanations; help seekers rarely contributed. Some help givers imitated the Socratic questioning used by their teachers to elicit understanding. This technique was unsuccessful in several instances because the questioners were at a loss as to how to continue the discussion when they received an unexpected response. Instead of continuing with another probe, the explanation providers gave up.

In the grade 7/8 sample there was no monitoring of the effectiveness of the explanations that were given. We could not find a single instance of an explanation giver testing to determine whether the explanation had been understood and appropriately used.

Discussion

Examination of specific interactions indicated a wide gap between the ideal model of explanation giving/receiving and student practice. The study had three main findings.

Students Infrequently Sought Explanations

Our investigation confirmed that students infrequently seek explanations, possibly because they did not know they needed help and did not know whom to ask if they did. Even though the correct answers to each exercise were distributed to students at the end of each group activity, students had difficulty interpreting differences between their work and the official answers. Graphs with apparently different data plots could express the same relationship because students were allowed to make their own decisions about the size and shape of the groups. The result was that students did not have good information on the performance of themselves or their peers. If students did not know they needed help, they would be unlikely to seek it (Markman, 1977, 1979).

The length of the treatment was also a factor in depressing explanation seeking. Constructing graphs was time consuming, especially for samples (a) and (b) who did everything by hand. The result was that students did not have a lot of data on which to make judgments about the competence of their peers. It also meant that they did not have enough interactions with group members to determine who would be willing to give help in a nonthreatening way, an important consideration as fear of ridicule is a major impediment to help seeking (Newman & Goldin, 1990; Newman & Schwager, 1993).

Students Received Inadequate Explanations

Many of the explanations were strewn with misconceptions. A reason is that they were as likely to be given by low- as high-ability students; neither rank in class nor specific competence (pretest scores on a correlational reasoning instrument) predicted explanation giving. In contrast, previous research on explanations has found that the more able give and the less able receive. Webb and Kenderski (1984) found that students with higher ability relative to other group members were more likely to give explanations, and Webb (1989) identified six studies in which giving explanations was correlated with overall ability.

Student certainty might be the key intervening variable here. For example, Miller and Brownwell (1975), in conservation studies involving young children, found that conservers were more willing to press their views than non-

conservers because they had more confidence in the validity of their claims. Tudge (1989, 1990) found that when children were asked to predict the outcome of a balance beam task, those with a decision rule that led to uncertainty in some conditions deferred to children who used a rule that produced certainty in all conditions, regardless of whether the rule was more or less advanced than their own. A critical element in the maintenance of the confidence of students with less powerful but more certain rules was Tudge's decision to give them no feedback on the accuracy of their predictions. Roth and Roychoudhury (1993) observed that senior high school students with greater conviction could persuade their peers, even if they were incorrect, if there was no other information source available. In a similar way, the lack of certainty about correlational reasoning performance described above might have increased the proportion of explanations offered by poorer performers.

Students' ability to give good explanations may have been further hampered by the absence of an appropriate language for describing correlational reasoning concepts. Basili and Sanford (1991) found that small-group discussions in junior college chemistry classes were impeded by students' lack of fluency in using scientific terms. The instructional treatments attempted to provide a few essential technical terms, but the students and teachers did not take to them, substituting fuzzy alternatives (e.g., referring to scale intervals as "places where you break it"). Although previous studies have found that nonverbal behavior is a central element in children's explanations (Bearison, 1982; Cooper, Ayers-Lopez, & Marquis, 1982; Forman, 1989; Mehan & Riel, 1982), it is desirable that it be accompanied by oral commentaries, even in highly visual tasks such as graph construction and interpretation. In giving demonstrations the explanation givers failed to describe their actions in conceptual terms so that the instructions could be generalized from one problem instance to another.

In addition to being flawed in content, the explanations that were offered by students suffered from many of the deficiencies of explanations exhibited by novice teachers. The student explanation givers used a teacher-as-teller strategy in which knowledge is transmitted rather than constructed, an approach characteristic of novice teachers (Florio-Ruane & Lensmire, 1990; Holt-Reynolds, 1992). Leinhardt and Greeno (1986) found that novice teachers were less proficient than experienced practitioners in using questioning to obtain information about student performance that could be used to redirect the lesson. The same difficulty was observed in our study with respect to explanation givers' inability to recover from an unusual help seeker response in a Socratic dialogue. Novice teachers lack the subject knowledge to identify suitable examples (Clermont, Borko, & Krajcik, 1994; Reynolds, 1992). The inability to think of alternate illustrations is the most likely reason why explanation givers in our investigations persisted with examples after the recipients had indicated they were not suitable. Like the students observed by Davis and Sigurdson (1992), our explainers were heavily dependent on the examples provided by teachers and continued with them even when it was clear that the teacher example had produced confusion in the minds of help seekers. There was also evidence that explanation givers reduced the demands of the curriculum by emphasizing completeness over other criteria and by

proposing to help seekers that they adopt simplistic problem definitions. Previous studies have found that these instructional procedures, used by experienced as well as novice teachers, reduce comprehension level tasks to algorithmic or recall assignments and have negative effects on conceptual understanding (Doyle & Carter, 1984; Miller, Leinhardt, & Zigmond, 1988; Sanford, 1985).

Another factor that may have reduced the effectiveness of explanations was help giver insensitivity to the learning needs of their peers. Mehan and Riel (1982) observed that when young children were explaining a game to their peers they usually started with an example provided by the explainer, whereas adult teachers were likely to begin with an example elicited from the child. In our investigations we found the same preference for beginning with the explainer's example, which was exacerbated by an unwillingness or inability to switch to the specific context in which the help seeker raised the problem. This practice placed the cognitive demands of translating the explanation into the desired context on students who may have been overloaded to begin with. Previous investigators (Cooper et al., 1982; Mehan & Riel, 1982) observed that student explainers used more directives than questions and were less likely than adults to obtain feedback on the learner's understanding as the explanation progressed, processes that we observed. Our student explainers employed a teaching strategy that was indifferent to the cognitive structures of learners, perhaps because they were unaware of the need to attend to the perspectives of others. It is also possible that explainer insensitivity to learner needs was a failure of caring. There were several instances of abusive verbal behavior in the recordings. Because previous studies have found that interpersonal behavior in cooperative learning groups changes over time (Ross & Raphael, 1990), it may be that explainer sensitivity might have increased if the experiments had continued over a longer period. Cooper et al. (1982) found that pairs of students who worked together frequently were more likely to share a teaching role.

Explanation Givers Did Not Monitor Help Seekers' Understanding of the Explanations

Hooper (1992) found that the only help giver response that influenced recipients' achievement was checking the helped person's understanding. Peterson and Swing (1985) observed that effective student groups were able to determine when explanations were appropriate and Webb (1989) suggested that student explanations were unlikely to have a beneficial impact unless they were understood and internalized. We observed few instances of help givers explicitly testing whether explanation seekers' needs had been satisfied. Although help givers occasionally inquired whether the recipient understood the explanation, more powerful tests such as asking the help seeker to recapitulate the explanation or apply it to another problem instance were very rare. The infrequency of such probes might be attributed to previously noted problems such as the small number of independent practice tasks assigned, uncertainty of explainers about their own performance, lack of awareness about the need to monitor understanding, and the failure of caring. Students' desire to complete the tasks quickly may also have reduced their willingness to monitor the impact of their explanations: speed of task completion is a criterion of group

success for many cooperative learning students (Holloway, 1990; Mulryan, 1992).

Conclusions

Implications for Teachers

Slavin (1987) argued that the developmental and motivational perspectives on cooperative learning could be reconciled if reward structures stimulated group members to help their peers. Such was not the case in this study. The poverty of explanation might explain why STAD was less effective than Group Investigation on higher-order achievement and more so on lower level objectives (Sharan, Kussell, Bejarano, Raviv, & Sharon, 1985). Two strategies for enhancing STAD come to mind. The first is to train students in interaction processes, as recommended by other cooperative learning approaches. The second is to select better examples for illustrating new ideas that are introduced in the teacher-directed phase of the lesson.

The training strategy might consist of the teacher giving explicit models of explanation in which the processes of asking for an explanation, giving one, and monitoring its impact could be modeled. Swing and Peterson (1982) found that the frequency and quality of explanations offered by primary age children could be enhanced through role playing that provides exemplars of good and poor explanations. Webb, Qi, Yan, Bushey, and Farivar (1990) found that role playing when combined with related strategies contributed to grade 7 students' ability to explain how to solve math problems. In another study we found that giving students feedback on their conversations had a positive impact. When students were given edited transcripts of their talk, along with teacher models of productive explanation, the frequency and quality of student interactions improved (Ross, in press). We also found that the teacher's interventions when students were working in groups influenced the explanations they gave to one another (Ross, 1994). The frequency and quality of students' explanations to their peers can also be enhanced through training in reciprocal questioning techniques in which students alternate between listener and critic (King, 1990; O'Donnell & Dansereau, 1992).

A different, but not mutually exclusive, strategy might consist of teachers selecting as concrete illustrations rich cases that optimize links between new ideas and existing understanding. The teachers in Davis and Sigurdson's (1992) study selected the examples used in whole-class demonstrations on the basis of their own out-of-school interests rather than on compatibility with student interests and experiences. Of particular concern are analogies that lend themselves to the development of misconceptions (Clement, Brown, & Zietsman, 1989, referred to these as "brittle anchors") because student explanations tend to be based on analogies to everyday experience (Hesse & Anderson, 1992). Because teacher illustrations are likely to be repeated by student explainers, contexts that create confusion, such as misleading analogies, are likely to persist in student discussions.

Finally, the failure of students to seek help when they need it might be reduced if teachers provided students with unambiguous appraisal criteria. Such criteria would enable those who need help to recognize it and to be able to identify peers with the academic ability to provide effective explanations.

Implications for Researchers

The findings from this research suggest several fruitful lines of investigation. The most interesting is to examine the impact of treatment duration on the effectiveness of explanation giving and receiving. Duration is likely to have an impact by increasing students' commitment to one another (although it should not be assumed that increased interaction will necessarily improve group cohesion) and by increasing students' familiarity with the skills and learning needs of their peers. Cooperative interactions over extended time periods might also result in explanation givers recognizing differences among students, becoming aware, for example, that not all students learn in the same way as oneself. Kagan (1992) concluded from a review of research on teacher growth that through practice teaching novice teachers reconstruct their images of teaching, and ultimately their practice, by obtaining knowledge of the variability of students. Increased duration might further impact explanation seeking by contributing to the development of a shared language, both verbal (Buckholdt & Wodarski, 1978) and nonverbal (Allan & Feldman, 1973).

Notes

1. Funding for this research was provided by the Ontario Ministry of Education through its Block Transfer Grant to the Ontario Institute for Studies in Education (#81-1118) and the Social Sciences and Humanities Research Council (#410-90-0834). The views expressed in the article do not necessarily reflect the views of the Ministry or SSHRC. Deborah Berrill contributed to the data collection in sample (a) and Anne Hogaboam-Gray contributed to the data collection in samples (b) and (c).
2. The software program was designed to support the correlational reasoning instruction. In addition to file management procedures, it contained simple commands to create multivariate scatterplots and calculate correlation coefficients (Hogaboam-Gray, Ross & Cousins, 1991).

References

Allan, V., & Feldman, R. (1973). Learning through tutoring: Low achieving children as tutors. *Journal of Experimental Education, 42*, 1-5.

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development, 59*(1), 87-96.

Bargh, J.A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology, 72*, 593-604.

Basili, P., & Sanford, J. (1991). Conceptual change strategies and cooperative group work in chemistry. *Journal of Research in Science Teaching, 28*, 293-304.

Bearison, D. (1982). New directions in studies of social interaction and cognitive growth. In F. Serafica (Ed.), *Social cognitive development in context* (pp. 199-221). New York: Guilford.

Brown, J., Collins, A., & Diguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*, 32-42.

Buckholdt, D., & Wodarski, J. (1978). The effects of different reinforcement systems on cooperative behaviors exhibited by children in classroom contexts. *Journal of Research and Development in Education, 12*, 50-68.

Clement, J., Brown, D., & Zietsman, A. (1989). Not all preconceptions are misconceptions: Finding "anchoring conceptions" for grounding instruction on students' intuitions. *International Journal of Science Education, 11*, 554-565.

Clermont, C., Borko, H., & Krajcik, J. (1994). Comparative study of the pedagogical content knowledge of experienced and novice chemical demonstrators. *Journal of Research in Science Teaching, 4*, 149-441.

Cooper, C., Ayers-Lopez, S., & Marquis, A. (1982). Children's discourse during peer learning in experimental and naturalistic situations. *Discourse Processes, 5*, 177-191.

Corno, L. (1989). Self-regulated learning: A volitional analysis. In B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement* (pp. 111-142). New York: Springer-Verlag.

Corno, L. (1992). Encouraging students to take responsibility for learning and performance. *Elementary School Journal, 93*(1), 69-83.

Cousins, J.B., & Ross, J. (1993). Improving higher order thinking skills "with" the computer: A comparative study. *Journal of Research on Computing in Education, 26*(1), 94-115.

Davis, B., & Sigurdson, S. (1992). *Talk in the mathematics classroom: Implementation and research concerns*. Paper presented at CSSE, Charlottetown.

Deering, P., & Meloth, M. (1990). *An analysis of the content and form of students' verbal interactions in cooperative groups*. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Doise, W., & Mugny, G. (1979). Individual and collective conflicts of centrations in cognitive development. *European Journal of Social Psychology, 9*, 105-108.

Doyle, W., & Carter, K. (1984). Academic tasks in classrooms. *Curriculum Inquiry, 14*, 129-149.

Florio-Ruane, S., & Lensmire, T. (1990). Transforming future teachers' ideas about writing instruction. *Journal of Curriculum Studies, 22*, 277-289.

Forman, E. (1989). The role of peer interaction in the social construction of mathematical knowledge. *International Journal of Educational Research, 13*, 55-70.

Good, T., Slavings, R., Harel, K., & Emerson, H. (1987). Student passivity: A study of question asking in K-12 classrooms. *Sociology of Education, 60*, 181-199.

Hertz-Lazarowitz, R. (1989). Cooperation and helping in the classroom: A contextual approach. *International Journal of Educational Research, 13*, 113-119.

Hesse, J., & Anderson, C. (1992). Students' conceptions of chemical change. *Journal of Research in Science Teaching, 29*, 277-299.

Hogaboam-Gray, A., Ross, J.A., & Cousins, J.B. (1991). *CORReoGRAPH—A program for solving correlational reasoning problems: User's guide*. Peterborough, ON: OISE Trent Valley Centre.

Holloway, S. (1990). *Contradictory meanings in teaching/learning "cooperation" via cooperative learning*. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Holt-Reynolds, D. (1992). Personal history-based beliefs as relevant prior knowledge in course work. *American Educational Research Journal, 29*, 325-349.

Hooper, S. (1992). Effects of peer interaction during computer-based mathematics instruction. *Journal of Educational Research, 85*, 180-189.

Johnson, D. (1990). *Reaching out: Interpersonal effectiveness and self-actualization* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Johnson, D., & Johnson, R. (1987). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Englewood Cliffs, NJ: Prentice-Hall.

Johnson, R., Johnson, D., & Stanne, M. (1986). Comparison of computer-assisted cooperative, competitive, and individualistic learning. *American Educational Research Journal, 23*, 382-392.

Kagan, D. (1992). Professional growth among preservice and beginning teachers. *Review of Educational Research, 62*, 129-169.

Kagan, S. (1988). *Cooperative learning: Resources for teachers*. Riverside, CA: University of California.

Karenbenick, S., & Knapp, J. (1988). Help seeking and the need for academic assistance. *Journal of Educational Psychology, 80*, 406-408.

Kempa, R., & Ayob, A. (1991). Learning interactions in group work in science. *International Journal of Science Education, 13*, 341-354.

King, A. (1989). Verbal interaction and problem-solving within computer-assisted cooperative learning groups. *Journal of Educational Computing Research, 5*(1), 1-15.

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal, 27*, 664-687.

Lawson, A., & Bealer, J. (1984). The acquisition of basic quantitative reasoning skills during adolescence: Learning or development. *Journal of Research in Science Teaching, 21*, 417-423.

Leinhardt, L., & Greeno, J. (1986). The cognitive skill of teaching. *Journal of Educational Psychology, 78*, 75-95.

Markman, E. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development, 48*, 986-992.

Markman, E. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development, 50*, 643-655.

Mehan, H., & Riel, M. M. (1982). Teachers' and students' instructional strategies. In L. Adler (Ed.), *Cross cultural research at issue* (pp. 171-188). New York: Academic Press.

Miller, S., & Brownwell, C. (1975). Peers, persuasion and Piaget: Dyadic interaction between conservers and nonconservers. *Child Development, 46*, 992-997.

Miller, S., Leinhardt, G., & Zigmond, N. (1988). Influencing engagement through accommodation: An ethnographic study of at-risk students. *American Educational Research Journal, 25*, 465-487.

Mulryan, C. (1992). Student passivity during cooperative small groups in mathematics. *Journal of Educational Research, 85*, 261-273.

Newman, R. (1990). Children's help-seeking in the classroom: The role of motivational factors and attitudes. *Journal of Educational Psychology, 82*, 71-80.

Newman, R., & Goldin, L. (1990). Children's reluctance to seek help with schoolwork. *Journal of Educational Psychology, 82*, 92-100.

Newman, R., & Schwager, M. (1993). Students' perceptions of the teacher and classmates in relation to reported help seeking in math class. *Elementary School Journal, 94*, 3-17.

O'Donnell, A., & Dansereau, D. (1992). Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 120-141) Cambridge, UK: Cambridge University Press.

Ontario Ministry of Education. (1988). *Intermediate and senior geography curriculum guidelines*. Toronto, ON: Queen's Park.

Ontario Ministry of Education. (1992). *Policy framework for the transition years*. Unpublished manuscript. Toronto, ON: Queen's Park.

Palincsar, A., & Brown, A. (1984). Reciprocal teaching of comprehension-monitoring activities. *Cognition and Instruction, 5*, 269-287.

Peterson, P.L., & Swing, S.R. (1985). Students' cognitions as mediators of the effectiveness of small-group learning. *Journal of Educational Psychology, 77*, 299-312.

Posner, G., Strike, K., Hewson, P., & Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*, 211-227.

Reynolds, A. (1992). What is competent beginning teaching? A review of the literature. *Review of Educational Research, 62*, 1-36.

Ross, J.A. (1994). *Explaining solution strategies: a case study of student talk in a math reform classroom.* Paper presented at the annual meeting of the Canadian Society for the Study of Education, Calgary.

Ross, J.A. (in press). Effects of feedback on student behavior in cooperative learning groups in a grade 7 math class. *Elementary School Journal.*

Ross, J.A., & Cousins, J.B. (1993a). Patterns of student growth in reasoning about multivariate correlational problems. *Journal of Educational Psychology, 85*, 1-17.

Ross, J.A., & Cousins, J.B. (1993b). Enhancing secondary school students' acquisition of correlational reasoning skill. *Research in Science and Technological Education, 11*(3), 191-206.

Ross, J.A., & Cousins, J.B. (in press). The impact of explanation seeking on student achievement and attitudes. *Journal of Educational Research.*

Ross, J.A., & Raphael, D. (1990). Communication and problem solving achievement in cooperative learning groups. *Journal of Curriculum Studies, 22*, 149-164.

Roth, W., & Roychoudhury, A. (1993). The concept map as a tool for the collaborative construction of knowledge: A micro-analysis of high school physics students. *Journal of Research in Science Teaching, 30*, 503-534.

Sanford, J. (1985). *Comprehension-level tasks in secondary classrooms. R&D Report No. 6199.* Austin, TX: R&D Centre for Teacher Education.

Sharan, S., Kussell, P.H.R., Bejarano, Y., Raviv, S., & Sharan, Y. (1985). Cooperative learning effects on ethnic relations and achievement in Israeli junior-high school classrooms. In R. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, C. Webb, & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 313-344). New York: Plenum Press.

Sharan, S., & Shachar, H. (1988). *Language and learning in the cooperative classroom.* New York: Springer Verlag.

Slavin, R. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin, 94*, 429-445.

Slavin, R. (1987). Developmental and motivational perspectives on cooperative learning: A reconciliation. *Child Development, 58*, 1161-1167.

Slavin, R. (1990). *Cooperative learning: Theory, research, and practice.* Englewood Cliffs, NJ: Prentice-Hall.

Slavin, R. (1992). When and why does cooperative learning increase achievement: Theoretical and empirical perspectives. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 145-173). Cambridge, UK: Cambridge University Press.

Solomon, J. (1992). The classroom discussion of science-based social issues presented on television: Knowledge, attitudes and values. *International Journal of Science Education, 14,* 431-444.

Swing, S., & Peterson, P. (1982). The relationship of student ability and small-group interaction to student achievement. *American Educational Research Journal, 19,* 259-274.

Tudge, J. (1989). When collaboration leads to regression: Some negative consequences of socio-cognitive conflict. *European Journal of Social Psychology, 19,* 123-189.

Tudge, J. (1990). *Collaborative problem solving in the zone of proximal development.* Paper presented at the annual meeting of the American Educational Research Association, Boston.

Webb, N. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13,* 21-39.

Webb, N. (1992). Testing a theoretical model of student interaction and learning in small groups. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 102-119), Cambridge, UK: Cambridge University Press.

Webb, N., & Kenderski, C. (1984). Student interaction and learning in small group and whole class settings. In P. Peterson, L. Wilkinson, & M. Hallinan (Eds.), *The social context of instruction: Group organization and group processes* (pp. 153-170). New York: Academic Press.

Webb, N., & Kenderski, C. (1985). Gender differences in small-group interaction and achievement in high- and low-achieving classes. In L. Wilkinson & C. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 209-236). Orlando, FL: Academic Press.

Webb, N., Qi, S., Yan, K., Bushey, B., & Farivar, S. (1990). *Cooperative small-group problem solving in middle school mathematics.* Paper presented at the American Educational Research Association, Boston.

Wilkinson, L., & Calculator, S. (1982). Requests and responses in peer-directed reading groups. *American Educational Research Journal, 19,* 107-120.

Wilkinson, L.C., & Spinelli, F. (1983). Using requests effectively in peer-directed instructional groups. *American Educational Research Journal, 20,* 479-501.

Yap, K., & Yeany, R. (1988). Validation of hierarchical relationships among Piagetian cognitive modes and integrated science process skills for different cognitive reasoning levels. *Journal of Research in Science Teaching, 25,* 247-281.

The Alberta Journal of Educational Research *Vol. XLI, No. 1, March 1995, 122-126*

Book Reviews

Genie: An Abused Child's Flight from Silence. Russ Rymer. New York: Harper Collins, 1993. 228 pages. ISBN: 0-06-016910-9.

Reviewed by *Tracey Derwing, University of Alberta*

This is the first book with a linguistics issue as a central theme that has ever moved me to tears. I was already familiar with much of the content, through an in-person account from one of the peripheral players involved, through Susan Curtiss's (1977) book on Genie's language development, and through Rymer's extensive article in the *New Yorker* on which this book is based. Even so, it is impossible to develop an immunity to the awful miscommunications on the part of people who had as their goal the rehabilitation of a child and the ramifications of the treatment she received. The story of Genie is a cautionary tale for researchers and educators; if ever there was a situation that exemplified the expression "good intentions pave the way to hell," this is it.

Genie was 13 years old when her mother took her and escaped the family home, where Genie's father had kept Genie captive, tied to a potty chair in a barren room for most of her life. The child was physically and cognitively underdeveloped; she had been spoken to only rarely and had not developed more than a few words (her productive vocabulary consisted of *stopit* and *nomore*). Rymer's book is a chronicle of what happened after Genie's plight was discovered in November of 1970 to the present. Along the way he discusses the nature of language acquisition, in particular the ongoing debate between those who believe that language is innate and those who maintain that it must be learned through interaction. Genie was thought to be an important piece of the puzzle: according to Chomsky's linguistic theory, the innate language acquisition device (latterly known as Universal Grammar) deteriorates over time, thus humans should not be able to learn a first language after the onset of puberty. If Genie succeeded in learning English, she would disprove the "critical period" hypothesis; she would also provide empiricists with strong evidence for the need for interaction in order to develop language.

The "Genie team," as Rymer dubbed them, was the group of doctors, psychologists, psychiatrists, linguists, and teachers who were to work with the girl. Rymer describes the first major meeting in which the scientists determined what the focus of their research with Genie should be. Interestingly, the evening before the conference, the Genie team and other experts viewed François Truffaut's film *The Wild Child* based on the diaries of physician Jean-Marc-Gaspard Itard who worked with Victor, another child who grew up in virtual isolation. Although Victor was discovered in 1800, the parallels with Genie are remarkable. The scientists were struck by the film; they shared the optimism that Itard's notes expressed with regard to Victor's progress. As Rymer indicates, however, Truffaut filmed only the first years of Itard's journals; Rymer deftly outlines the similarities between Genie and Victor later on in these two

natural experiments—Victor's slide from celebrity to obscurity, from constant testing and attention to benign neglect. Although there were experts at the Genie conference who expressed some concerns about the direction the research was to take, namely, Genie's linguistic development, those most closely involved were convinced that what was in the best interests of research was also in the best interests of the child.

Rymer extensively interviewed as many of the principals as possible; he also obtained court records, personal letters, and other documentation. He immersed himself in linguistic theory (both sides of the fence) in order to present the whole story as fairly as he could. The only person with whom he had no contact (either direct or indirect) is Genie herself, yet she is the heart and soul of the book. Nearly all undergraduate linguistics students know that Genie was a most compelling communicator; they have heard anecdotes about passersby on the street handing her a dollie or a necklace or some other object—all reporting that the little girl had somehow let them know that she wanted it. Susan Curtiss, who spent years with Genie, initially tracing her language development but eventually becoming her friend and supporter, is quoted as saying "Genie is the most powerful, most inspiring person I have ever met. I'd give up my job, I'd change careers to see her again" (p. 221). To Rymer's credit, Genie has communicated here as well.

Rymer outlines the events surrounding Genie's initial stay at Children's Hospital of Los Angeles, her placement in the home of her teacher, her four-year stay in the home of her psychologist, her short reunification with her mother, and a series of foster homes after that. There were misunderstandings and disagreements from the start as to what Genie needed and who would best provide for her needs. Disagreement turned to acrimony and vendetta; a confounding of roles of researcher and foster parent led to disaster. The two individuals who are most perplexing in the sordid details of the book are Jean Butler Ruch and David Rigler. Jean Butler was Genie's teacher; she clearly loved Genie and wanted to keep her in her home. Although she was officially a part of the Genie team, evidently Butler was not a team player. She had had a record of difficulty working with others, and she evidenced similar problems in her role at Children's Hospital. When she was denied foster parent status by the Department of Public Social Services, she started a campaign against those who were associated with the Genie team, writing letters to other scientists, castigating the team's efforts. She also befriended Genie's biological mother and eventually convinced her to sue nearly everyone involved in the case. Rigler, Genie's psychologist, seems to have been largely motivated by academic ambition and greed, yet he and his family sacrificed a great deal by having Genie in their home for four years, where she was apparently making good progress. Within a month of hearing of a failure to receive renewal on an NIMH grant, however, Rigler sent Genie back to her biological mother.

In an interview with Jay Shurley, a psychiatrist who was a consultant from the outset on Genie's case and one of the only people who is still allowed to see Genie, it is revealed that she is now in an institution. "The way I think of Genie, she was this isolated person, incarcerated for all those years, and then she emerged and lived in a more reasonable world for a while, and responded to

this world, and then the door was shut and she withdrew again and her soul was sick" (p. 213).

The shutting of the door is the crucial point of this book. Although all involved were convinced of the ethical nature of their research, it is the bungling of ethical issues, the failure to see Genie as a whole, a person who will no doubt live for some years to come, that resulted in her regression. The progress that was so carefully documented in test after test after test, and years of intense observation has been lost. There was no long-term plan for Genie, no integrated approach to her rehabilitation. The expectations of her were too high—she was to solve a question that has not been resolved for centuries; when she didn't provide an answer, she was shunted off to an institution through a litigious and bureaucratic route.

What Rymer has provided in this book is a complete picture of the complicated mess that was made of Genie's treatment and the very limited knowledge gleaned about language acquisition as a result of her study. Although some linguists claim that Genie's failure to acquire grammar supports the innatist stand, interactionists such as Catherine Snow ask, "How could a child who lacked language because she had been shut away from her mother be proof of the contention that our mothers don't teach us language?" (p. 157).

The book has a few oversimplifications with regard to linguistic matters that are mildly irritating, but none that affect the basic thesis. It is unfortunate that no references are listed. This is intended to be a book for the layperson; however, the author owes the reader original source information. These complaints pale in light of the overall impression. The highly readable style, the careful documentation of detail, and the fascinating nature of the subject matter make this a thought-provoking and thoroughly engaging account of a tragedy.

Reference

Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern-day "wild child."* New York: Academic Press.

Tracey Derwing is an associate professor who teaches teaching English as a second language. Her research interests include native speaker-nonnative speaker discourse, adult second language literacy, and citizenship education for immigrants.

Violence and Abuse in the Lives of People With Disabilities. Dick Sobsey. Baltimore: Paul H. Brookes, 1994.

Reviewed by *Gregor Wolbring, German Council of Self-Determined Living Centres*

I am pleased to review the book *Violence and Abuse in the Lives of People With Disabilities* by Dick Sobsey. In order to place the tone of my review in context, I should state that I am not from the educational field by profession; I am a biochemist. Nonetheless, as a disability activist, I feel confident to review this book.

Although several books are available on the topic of abuse and disabilities (e.g., Garbarino, Brookhouser, & Authier, 1987; Morgan, 1987; Senn, 1987), Sobsey's 1994 book is the first to provide a comprehensive model that helps us understand abuse and other forms of maltreatment (e.g., euthanasia has previously been seen as an independent phenomenon).

As a disability activist, I feel that more should be written about the violence against people with disabilities. Drawing on previous work (Sobsey, Gray, Wells, Pyper, & Reimer-Heck, 1991; Sullivan, Brookhauser, Scanlan, Knutson, & Schulte, 1991; Turk & Brown, 1992; Westcott, 1993), Sobsey outlines in a clear and detailed fashion the probability that a disabled person will be abused by a family member (16.5%), an acquaintance or neighbor (16.5%), a disability service provider (28%), or a stranger (6.6%). In addition the book clearly describes how the risk increases that people with a mental disabilities will be victimized. Using data from Wilson and Brewer (1992), Sobsey suggests that mentally disabled people are more than 10 times as likely as other citizens to be victims of robbery or sexual assault.

Violence and Abuse in the Lives of People With Disabilities shows us how society trivializes or decriminalizes offenses against disabled people. For example, rather than describe an act of murder of a disabled person as murder, we use terms such as euthanasia or assisted suicide (for further discussion of this issue, see Luckasson, 1992).

The tables and statistics in the book are powerful reminders of the frequency and severity of violence against people with disabilities that should make readers shudder with disgust. *Violence and Abuse in the Lives of People With Disabilities* will surely generate greater demand for laws and standards that protect and integrate disabled people. Equally powerful is the discussion of Nazi history and the presentation of an ecological model of abuse.

Perhaps the most compelling part of the book is the elaborate section (nearly half the book) that describes how abuse can be prevented. This section is divided into several topics: Empowering individuals to resist abuse; Families and other caregivers: support and selection; Building safer environments; Law and law enforcement; Changing attitudes that disinhibit violence; Healing the consequences of abuse; and Prevention and intervention teams. I would like to single out the topic "Changing attitudes that disinhibit violence" for discussion in this review. This section clearly shows the "Catch 22" situation that people with disabilities face. Society, especially through the media, presents pictures of disabled people that can best be understood as two stereotypes. Either disabled people are portrayed as suffering individuals who are a burden to society or they are seen as superheroes. The stereotype of the suffering individual who is a burden to society is used a lot these days in an attempt to push for genetic testing, assisted suicide, or euthanasia ("mercy killing" is the phrase often used here). The stereotype of superhero portrays the disabled person as productive against all odds. Both stereotypes lead to a situation where people with a disability are viewed and treated differently than others. Being treated differently leads disabled people to behave differently; it also threatens their self-esteem. Their own different behavior often acts to confirm the stereotypes that people hold. Having said this, I feel that the part of

Sobsey's book that deals with attitudes toward people with a disability and how attitudes might be changed is of utmost importance.

To summarize, *Violence in the Lives of People With Disabilities* extends previous work with an integrated ecological model of violence that includes elements from social labeling theory and ambivalence theory. From this perspective, child abuse and sexual assault can no longer be seen as unique phenomena, but rather as part of the larger spectrum of negative outcomes that result from discriminatory attitudes.

References

Garbarino, J., Brookhouser, P. E., & Authier, K.J. (Eds.). (1987). *Special children-special risks: The maltreatment of children with disabilities*. New York: Aldine de Gruyter.

Luckasson, R. (1992). People with mental retardation as victims of crime. In R.W. Conley, R. Luckasson, & G.N. Bouthilet (Eds.), *The criminal justice system and mental retardation: Defendants and victims* (pp. 209-220). Baltimore, MD: Brookes.

Morgan, S.R. (1987). *Abuse and neglect of handicapped children*. Boston, MA: College Hill Press.

Senn, C.Y. (1987). *Vulnerable: Sexual abuse and people with an intellectual handicap*. Downsview, ON: G. Allan Roeher Institute. (ERIC Document Reproduction Service No. ED 302 975)

Sobsey, D., Gray, S., Wells, D., Pyper, D., & Reimer-Heck, B. (1991). *Disability, sexuality, and abuse: An annotated bibliography*. Baltimore, MD: Brookes.

Sullivan, P.M., Brookhauser, P.E., Scanlan, J.M., Knutson, J.F., & Schulte, L.E. (1991). Pattern of physical and sexual abuse of communicatively handicapped children. *Annals of Otology, Rhinology, and Laryngology, 100*(3), 188-194.

Turk, V., & Brown, H. (1992, August 5-7). *Sexual abuse of adults with learning disabilities*. Paper presented at the Conference of the International Association for Scientific Study of Mental Deficiency, Brisbane, Australia.

Westcott, H. (1993). *Abuse of children and adults with disabilities*. London: National Society for Prevention of Cruelty to Children.

Wilson, C., & Brewer, N. (1992). The incidence of criminal victimization of individuals with an intellectual disability. *Australian Psychologist, 2*, 114-117.

Gregor Wolbring is a research scientist in the biochemistry department at the University of Calgary. As a disability activist he has delivered several invited presentations on disability-related topics, genetics, reproductive technologies, ethics, assisted suicide, euthanasia, and the interaction between the women's movement and the disability movement.

For Contributors

Please prepare manuscripts in accordance with these guidelines.

Format

Manuscript in triplicate.
Typewritten or printed, double-spaced throughout.
Abstract of approximately 100 words on a separate sheet.
Manuscript not to exceed 6,500 words excluding graphics.
Author's name and affiliation on a separate title page; title only on first page of manuscript to ensure anonymity in the review process.

Referencing

Sources in parentheses after each reference (direct or otherwise), giving author's name (unless mentioned in text), year of publication, and page number(s) for direct quotes.
Direct quotes of 40 words or less in double quotation marks incorporated into text; those longer than 40 words in indented block format. Page numbers must be given.
All sources listed alphabetically at the end of the manuscript under the heading *References* using the APA style.
Explanatory notes, numbered consecutively and marked in the text with superscript numerals, may appear before the References under the heading *Notes*; any citations in notes follow the same format as other references.

Graphics

Each table and figure on a separate page at the end of the manuscript.
Figures, charts, and diagrams in camera-ready form.
Number tables in Arabic numerals and center over each heading, e.g., *Table 1*.
Number figures in Arabic numerals before the caption at foot of each graph, e.g., *Figure 1. ...*
Indicate placement of figures and tables in text, i.e., *Insert Table 3 here.*

Style

AJER's editorial style conforms closely to the *Publication Manual of the American Psychological Association*.
Editorial changes may be made to manuscripts.
For spelling, consult *Webster's New Collegiate Dictionary;* spelling in quoted material remains as in the original.

Diskettes

If manuscript is accepted for publication please supply a diskette from a Macintosh, IBM, or IBM compatible system. Microsoft Word and WordPerfect are the preferred software.

Submission to *AJER* implies the manuscript is not being considered for publication elsewhere.

AJER is indexed in the *Alberta Education Index, Canadian Education Index, Current Contents/Social and Behavioral Sciences, ERIC/Current Index to Journals in Education,* and *Social Science Citation Index;* appropriate articles are abstracted in *Cognitive Development Abstracts, Educational Administration Abstracts, Educational Technology Abstracts, Family Abstracts, Language Behavior Abstracts, Multicultural Education Abstracts, Psychological Abstracts, Research into Higher Education Abstracts, School Organization and Management Abstracts,* and *Sociology of Education Abstracts,* and reviewed by the *Journal for Research in Mathematics Education.*

If undelivered return to:

The Alberta Journal of Educational Research
845 Education Centre South
University of Alberta
Edmonton, AB
T6G 2G5

Publications Mail Registration Number 1438